

```

1:  gb_ba:*
2:  gb_htg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_nam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query	Length	DB	ID	Description
No.						
1	1192	100.0	1192	6	AX431088	AX431088 Sequence
2	1082.4	90.8	1125	9	BC013587	BC013587 Homo sapi
3	1012	87.9	1049	9	BD191093	BD191093 186 Human
4	808.8	64.9	165110	9	AL450311	AL450311 Human DNA
5	801.6	67.2	173341	2	AC021954	AC021954 Homo sapi
6	473	39.7	483	6	AR431227	AR431227 Sequence
7	473	39.7	483	6	BD108780	BD108780 EST and e
8	452.6	38.0	550	6	AX431082	AX431082 Sequence
9	432.8	36.3	643	6	AR431427	AR431427 Sequence
10	351.8	32.9	519	6	BD110980	BD110980 EST and e
11	391.8	32.9	519	6	BD110980	BD110980 EST and e
12	375.4	31.5	480	11	G22651	G22651 Human STS W
13	307.4	25.8	1538	11	BC024943	BC024943 Mus muscu
14	306	25.7	275924	2	AC111623	AC111623 Rattus no
15	294	24.7	240657	2	AC111675	AC111675 Rattus no
16	250.8	21.0	258815	2	AC127817	AC127817 Rattus no
17	250.6	21.0	215050	2	AC127417	AC127417 Mus muscu
18	235.4	19.8	138070	2	AC109783	AC109783 Mus muscu
19	235.8	19.5	364	6	AR425705	AR425705 Sequence
20	233.8	19.5	364	6	BD121258	BD121258 EST and e
21	88.2	7.4	238937	2	AC094926	AC094926 Rattus no
22	75.4	6.3	177226	10	AC121858	AC121858 Mus muscu
23	75.4	6.3	183975	10	AL590988	AL590988 Mouse DNA
24	75	6.3	986	6	AX644366	AX644366 Sequence
25	69.6	5.8	577	6	AX535036	AX535036 Sequence
26	65.2	5.5	737	6	AX644276	AX644276 Sequence
27	64	5.4	666	6	AX431080	AX431080 Sequence
28	63.6	5.3	626	6	AX431141	AX431141 Sequence
29	62.2	5.2	879	6	AX644313	AX644313 Sequence
30	60.4	5.1	1545	6	AX535007	AX535007 Sequence
31	59.8	5.0	1623	6	AX497809	AX497809 Sequence
32	58.4	4.9	510	6	AX431074	AX431074 Sequence
33	57.8	4.8	2910	6	AX431097	AX431097 Sequence
34	57.6	4.8	110000	2	PRM46P1_10	Continuation (11 o
35	57.4	4.8	254733	2	AC1117075	AC1117075 Dictyoste
36	57.2	4.8	27785	2	AC1116978	AC116978 Dictyoste
37	57.2	4.8	183648	3	AC1117076	AC1117076 Dictyoste
38	57.2	4.8	250743	3	AE014836	AE014836 Plasmodu
39	57	4.8	663	6	AX431086	AX431086 Sequence
40	56.8	4.8	200846	2	AC146514	AC146514 Pan trogl
41	56.6	4.7	250029	3	AE014839	AE014839 Plasmodu
42	56.4	4.7	1066	10	RRHYDSTL	X63410 R.rattus mR
43	56.4	4.7	256879	3	AC115682	AC115682 Dictyoste
44	55.6	4.7	57538	3	AC115682	AC115682 Dictyoste
45	55.6	4.7	182871	3	AC111716	AC111716 Dictyoste

## ALIGNMENTS

RESULT 1					
AX431088			1192 bp.	DNA	linear
LOCUS	AX431088				
DEFINITION	Sequence 31 from Patent WO240535.				
ACCESSION	AX431088				
VERSION	AX431088.1				
KEYWORDS					
SOURCE	Homio sapiens (human)				
ORGANISM	Homio sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 Recipon, H., Liu, C., Salceda, S., Sun, Y., Cafferty, R. and				
AUTHORS	Macinn, R.A.				
TITLE	Compositions and methods relating to ovary specific genes and				

841 CAGCTGGATGCCCAACATGTTCCGATGCTGTGGAGACATGCCGAGTCTCTCTGCTTA 500

center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcdpax1.stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
Series: IRAC Plate: 14 Row: 1 Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21687152.

## FEATURES

source

1. 1129

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:9596 IMAGE:3896656"

/tissue\_type="pancreas, epithelioid carcinoma"

/clone\_id="NIH\_MGC\_70"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 1129

/gene="C10orf35"

/db\_xref="LocusID:219738"

160. 525

/codon\_start=1

/product="chromosome 10 open reading frame 35"

/protein\_id="AAH13587.1"

/db\_xref="GI:15488920"

/db\_xref="LocusID:219738"

/translation="VRIILANGBIIVDDPRVTTTPPRGSIIPRGSFNRHGADPG  
GPGRQQAGARLGAASPPNDLNRQLVNMGPQWHLGNHVAVPVTSILLFLMLWG  
VRGLLVGLVYLVSHLSOR"

## ORIGIN

Query Match 90.8%; Score 1082.4; DB 9; Length 1129;  
Best Local Similarity 99.9%; Pred. No. 6e-261; Indels 0; Gaps 0;

Matches 1083; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 2 AACCGCTGAGCCGCCGAGTGAAGCGCCGAGGCCGAGAGTGGCGCTGCAGACAC 61
Db 1 AACCGCTGAGCCGCCGAGTGAAGCGCCGAGGCCGAGAGTGGCGCTGCAGACAC 60

QY 62 AGCATCTACAGAGTGGTCACTCTGTGAATCATCTACATGCAAGCTCCCTCAATT 121
Db 61 AGCATCTACAGAGTGGTCACTCTGTGAATCATCTACATGCAAGCTCCCTCAATT 120

QY 122 TCTGTGACAGCCCATCAAGGACCAAGCCGCTGGAGAGATGGTGGGATCTTGGCAAT 181
Db 121 TCTGTGACAGCCCATCAAGGACCAAGCCGCTGGAGAGATGGTGGGATCTTGGCAAT 180

QY 182 GGGGAATCTGTGACGACGACGCCCGAGTGAAGCACTAACCCAGCCACCAAGAGT 241
Db 181 GGGGAATCTGTGACGACGACGCCCGAGTGAAGCACTAACCCAGCCACCAAGAGT 240

QY 242 AGCATCTGACAGAGCTTCTTCATATAGGGGCAATGTGCTCCCGAGGGGTCCGGC 301
Db 241 AGCATCTGACAGAGCTTCTTCATATAGGGGCAATGTGCTCCCGAGGGGTCCGGC 300

QY 302 CCCCGCAGAGCAGGAGGTGCAGGCTGGTGTCTCAATGCCCTTCAATGAAGCTC 361
Db 301 CCCCGCAGAGCAGGAGGTGCAGGCTGGTGTCTCAATGCCCTTCAATGAAGCTC 360

QY 362 AACCGGAGAGTGTGAACATGGGCTTCCGCAATGGATCTGGGCAACCATGCTGGAG 421
Db 361 AACCGGAGAGTGTGAACATGGGCTTCCGCAATGGATCTGGGCAACCATGCTGGAG 420

QY 422 CCGGTGACCTCATCTGCTCTCTTCTGCTCATGATGTTGGTGTGGTGGCTCCGC 481
Db 421 CCGGTGACCTCATCTGCTCTCTTCTGCTCATGATGTTGGTGTGGTGGCTCCGC 480

QY 482 CTGGTGGCTTGTCTACCTGTGTCCCACTGATGACGGGTGACCTTGAAGGCTGAT 541
Db 481 CTGGTGGCTTGTCTACCTGTGTCCCACTGATGACGGGTGACCTTGAAGGCTGAT 540

```

```

QY 542 AGGGGTGGGTTTGTGAGAGGGAATTGCGGGCTTGGTGTGAGAGGACATATTTGA 601
Db 541 AGGGGTGGGTTTGTGAGAGGGAATTGCGGGCTTGGTGTGAGAGGACATATTTGA 600

QY 602 GGGATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 661
Db 601 GGGATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660

QY 662 GTTGTGTAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 721
Db 661 GTTGTGTAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720

QY 722 GTCCGTGTGTCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 781
Db 721 GTCCGTGTGTCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 780

QY 782 CAGAGCAAGAAAGCAGATGATGATGATGATGATGATGATGATGATGATGATG 841
Db 781 CAGAGCAAGAAAGCAGATGATGATGATGATGATGATGATGATGATGATGATG 840

QY 842 AGCTGATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 901
Db 841 AGCTGATGACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 900

QY 902 GGAGCAGACTTGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 961
Db 901 GGAGCAGACTTGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960

QY 962 TGTTTAATGATATTTTATGAGCTATTTATGATGATGATGATGATGATGAT 1021
Db 961 TGTTTAATGATATTTTATGAGCTATTTATGATGATGATGATGATGATGAT 1020

QY 1022 CTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1081
Db 1021 CTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080

QY 1082 AAAA 1085
Db 1081 AAAA 1084

RESULT 3
BD191093
LOCUS 186 human secreted proteins. 1049 bp DNA linear PAT 17-JUL-2003
ACCESSION BD191093
VERSION BD191093.1 GI:33000832
KEYWORDS JP 2002510192-A/57.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1049)
AUTHORS Ruben,S.M., Rosen,C.A., Fischer,C.L., Soppet,D.R., Carter,K.C.,
Bednarik,D.P., Endress,G.A., Yu,G.L., N.J., Feng,P., Young,P.E.,
Greene,J.M., Fertle,A.M., Duan,R., Hu,J.S., Florence,K.A.,
Olson,H.S., Eder,R., Brewer,L.A., Moore,P.A., Shi,Y.,
Lafleur,D.W., Li,Y., Zeng,Z. and Kyaw,H.
186 human secreted proteins
Patent: JP 2002510192-A 57 02-APR-2002;
HUMAN GENOME SCIENCES INC
PN JP 2002510192-A/57
PD 02-APR-2002
PF 06-MAR-1998 JP 199853883
PR 07-MAR-1997 US 60/040162, 07-MAR-1997 US 60/040333 PR
07-MAR-1997 US 60/038621, 07-MAR-1997 US 60/040161 PR
07-MAR-1997 US 60/040626, 07-MAR-1997 US 60/040334 PR
07-MAR-1997 US 60/040336, 07-MAR-1997 US 60/040163 PR
11-APR-1997 US 60/043580, 11-APR-1997 US 60/043568 PI
M RUBEN, CRAIG A ROSEN, CARLE L FISCHER, DANIEL R SOPPET,
KENNETH C CARTER, DANIEL P BEDNARIK, GREGORY
A ENDRESS, GUO LIANG
PI YU, JIAN NI,
PI PING FENG, PAUL E YOUNG, JOHN M GREENE, ANN

```

M FERRIE, ROXANNE DUAN,  
 PI JING SHAN HU, KIMBERLY A FLORENCE, HENRIK  
 S OLSEN, REINHARD EBERER,  
 PI LAURIE A BREWER, PAUL A MOORE, YANGSU SHI, DAVID W LAFLEUR PI  
 , YI LI, ZHI ZHEN ZENG,  
 PI HLA KYAM  
 G01N33/50,  
 C12N5/12, C12N5/10, C12N1/21, C07K14/47, C07K16/18, C12Q1/68, PC  
 G01N33/53, G01N33/68, A61K38/17  
 PC  
 CC Strandedness: Double;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers.

ORIGIN  
 Query Match 84.9%; Score 1012; DB 6; Length 1049;  
 Best Local Similarity 99.3%; Pred. No. 3e-243;  
 Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

46 TCGGCGCTGCGACACAGCATCTACTCAGCGTGGGTCACTCTGTGTGAACATCTGACTG 105  
 1 TCGGCGCTGCGACACAGCATCTACTCAGCGTGGGTCACTCTGTGTGAACATCTGACTG 60  
 106 CAAGCTCTCCCTCAATTTCTGTGTGAGCCCATCAGAGGACCAAGCGCTG3GAGATGT 165  
 61 CAAGCTCTCCCTCAATTTCTGTGTGAGCCCATCAGAGGACCAAGCGCTG3GAGATGT 120  
 166 GCGATCTTGGCCCAATGGGAAATCGTGCAGAGACGAGACCCCGAGTGGAGCACTAC 180  
 121 GCGATCTTGGCCCAATGGGAAATCGTGCAGAGACGAGACCCCGAGTGGAGCACTAC 180  
 226 CGAGGCACCAAGAGGTGACATTCCTCGACAGAGCTTCTCAATGAGGAGCATGCTGC 285  
 181 CGAGGCACCAAGAGGTGACATTCCTCGACAGAGCTTCTCAATGAGGAGCATGCTGC 240  
 286 CCGAAGGGGCTCTGCGCCCGCCGACAGACAGGAGGAGGAGGAGGAGGAGGAGGAGG 345  
 241 CCGAAGGGGCTCTGCGCCCGCCGACAGACAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
 346 CCGCTTCAATAGCTCAACCGGAGCGTGTGAGATGAGGAGGAGGAGGAGGAGGAGGAGG 405  
 301 CCGCTTCAATAGCTCAACCGGAGCGTGTGAGATGAGGAGGAGGAGGAGGAGGAGGAGG 360  
 406 CAACCATCTGTGAGAGCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 465  
 361 CAACCATCTGTGAGAGCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
 466 TGTTCGTGGCT 525  
 421 TGTTCGTGGCT 480  
 526 ACGCTGAGGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 585  
 481 ACGCTGAGGAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540  
 586 AGCAGGCAATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 645  
 541 AGCAGGCAATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
 646 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 703  
 601 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
 704 TTTCCAAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 763  
 661 TTTCCAAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
 764 CATTAGTGGCAAAATCGAGCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 823

Db 721 CATTAGTGGCAAAATCGAGCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
 QY 824 TTAGCCGAGAGAGATCTCAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 883  
 Db 781 TTAGCCGAGAGAGATCTCAGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
 QY 884 GAGCT 943  
 Db 841 GAGCT 900  
 QY 944 CTTTCTTGAAGCTGTTTGTGTTTGAAGCTGTTTGTGTTTGAAGCTGTTTGTGTTTGAAG 1003  
 Db 901 CTTTCTTGAAGCTGTTTGTGTTTGAAGCTGTTTGTGTTTGAAGCTGTTTGTGTTTGAAG 960  
 QY 1004 ACAGTGTGAGTGGCATCTTATTTATTTGAGACCTTTCAATTAATTAATTAATTAAT 1063  
 Db 961 ACAGTGTGAGTGGCATCTTATTTATTTGAGACCTTTCAATTAATTAATTAATTAAT 1020  
 QY 1064 AAAAAAAAAAAAAA 1077  
 Db 1021 AAAAAAAAAAAAAA 1034

RESULT 4  
 AL450311  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 165110)  
 Direct Submission  
 Submitted (12-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: hinxton@sanger.ac.uk  
 Requesters: clonerequest@sanger.ac.uk  
 On Jul 8, 2001 this sequence version replaced gi:14575291.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Emi, EMBL; Sw,  
 SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP  
 database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 10, constructed by the Sanger Centre Chromosome 10  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr10  
 RPL1-343J3 is from the library RPL1-11.2 constructed by the group  
 of Pletier de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm

VECTOR: pBACe3.6  
 This sequence is the entire insert of clone RPL1-343J3 The true  
 left end of clone RPL1-242620 is at 139955 in this sequence. The  
 true right end of clone RPL1-404C6 is at 6588 in this sequence.  
 Location/Qualifiers  
 1. 165110



/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /clone="RP11-343J3"  
 /clone\_lib="RPC1-11.2"  
 7. .147  
 repeat\_region /note="THE1C repeat: matches 2. .142 of consensus"  
 151. .293  
 repeat\_region /note="MIR repeat: matches 2. .148 of consensus"  
 1136. .1187  
 /note="26 copies 2 mer gt 98% conserved"  
 repeat\_region 2960. .3272  
 /note="AluSg repeat: matches 1. .310 of consensus"  
 4144. .4270  
 repeat\_region /note="AluSx repeat: matches 1. .134 of consensus"  
 4319. .4494  
 /note="AluSx repeat: matches 118. .293 of consensus"  
 5662. .6438  
 repeat\_region /note="L1MC repeat: matches 272. .1095 of consensus"  
 6916. .8080  
 repeat\_region /note="L1MC repeat: matches 1168. .2367 of consensus"  
 9934. .10087  
 repeat\_region /note="MIR repeat: matches 25. .184 of consensus"  
 10353. .10481  
 repeat\_region /note="MIR repeat: matches 65. .194 of consensus"  
 11077. .11372  
 /note="AluSx repeat: matches 1. .294 of consensus"  
 11600. .11639  
 repeat\_region /note="20 copies 2 mer tc 95% conserved"  
 12479. .12608  
 repeat\_region /note="65 copies 2 mer at 90% conserved"  
 13433. .13876  
 repeat\_region /note="MTRIC repeat: matches 1. .466 of consensus"  
 14372. .14576  
 repeat\_region /note="AluSg repeat: matches 1. .301 of consensus"  
 14577. .14633  
 /note="AluSg repeat: matches 243. .299 of consensus"  
 14843. .14943  
 repeat\_region /note="MER86 repeat: matches 9. .109 of consensus"  
 15404. .15613  
 repeat\_region /note="MIR repeat: matches 3. .213 of consensus"  
 16626. .16705  
 repeat\_region /note="MIR repeat: matches 60. .140 of consensus"  
 16928. .16993  
 /note="33 copies 2 mer gg 66% conserved"  
 17476. .17569  
 repeat\_region /note="MER81 repeat: matches 2. .114 of consensus"  
 17719. .18069  
 repeat\_region /note="L1MC4 repeat: matches 7617. .7977 of consensus"  
 18117. .18345  
 /note="MIR repeat: matches 7. .262 of consensus"  
 19169. .19400  
 repeat\_region /note="MIR repeat: matches 5. .239 of consensus"  
 19436. .19627  
 repeat\_region /note="L1MC5 repeat: matches 7728. .7917 of consensus"  
 19642. .19716  
 /note="MIR repeat: matches 48. .131 of consensus"  
 19969. .20010  
 repeat\_region /note="21 copies 2 mer tg 100% conserved"  
 20623. .20704  
 repeat\_region /note="41 copies 2 mer gt 85% conserved"  
 20738. .20773  
 repeat\_region /note="U2 repeat: matches 1. .36 of consensus"  
 20802. .20863  
 /note="MTR1F repeat: matches 1. .62 of consensus"  
 20889. .21263  
 repeat\_region /note="MTR1F repeat: matches 188. .541 of consensus"  
 21463. .21618  
 repeat\_region /note="MIR repeat: matches 46. .192 of consensus"  
 22019. .22326  
 /note="AluB repeat: matches 1. .306 of consensus"  
 22381. .22564

repeat\_region /note="U2 repeat: matches 2453. .2629 of consensus"  
 22896. .23174  
 repeat\_region /note="MTR1F repeat: matches 117. .413 of consensus"  
 23215. .23346  
 /note="MIR repeat: matches 48. .188 of consensus"  
 23388. .23532  
 /note="U2 repeat: matches 2097. .2230 of consensus"  
 23533. .23837  
 repeat\_region /note="AluB repeat: matches 1. .303 of consensus"  
 23838. .24137  
 repeat\_region /note="U2 repeat: matches 1754. .2097 of consensus"  
 24291. .24581  
 /note="AluSx repeat: matches 1. .300 of consensus"  
 24653. .24850  
 repeat\_region /note="MIR repeat: matches 1. .200 of consensus"  
 26189. .26685  
 /note="CPG island"  
 /evidence=not\_experimental  
 29453. .30918  
 /note="CPG island"  
 /evidence=not\_experimental  
 31084. .31156  
 repeat\_region /note="MIR repeat: matches 65. .138 of consensus"  
 31618. .31876  
 /note="AluB repeat: matches 29. .275 of consensus"  
 32767. .32830  
 repeat\_region /note="MIR repeat: matches 76. .139 of consensus"  
 33050. .33178  
 /note="43 copies 3 mer tcc 72% conserved"  
 35112. .35201  
 /note="CPG island"  
 /evidence=not\_experimental  
 36551. .36604  
 /note="27 copies 2 mer ac 94% conserved"  
 37585. .38254  
 /note="L1PA15 repeat: matches 5480. .6157 of consensus"  
 38265. .38561  
 /note="AluSx repeat: matches 1. .300 of consensus"  
 39214. .39423  
 repeat\_region /note="MIR repeat: matches 3. .219 of consensus"  
 39489. .39801  
 /note="AluSx repeat: matches 1. .292 of consensus"  
 40050. .40189  
 repeat\_region /note="MIR repeat: matches 1. .144 of consensus"  
 41057. .41290  
 /note="MIR repeat: matches 8. .240 of consensus"  
 41296. .41460  
 repeat\_region /note="U2 repeat: matches 2569. .2730 of consensus"  
 41506. .41944  
 /note="U2 repeat: matches 1916. .2416 of consensus"  
 42388. .42698  
 /note="U2 repeat: matches 1448. .1779 of consensus"  
 44193. .44579  
 /note="THE1C repeat: matches 1. .371 of consensus"  
 44600. .44956  
 repeat\_region /note="MTR16A repeat: matches 90. .445 of consensus"  
 45240. .45300  
 /note="MER8A repeat: matches 37. .97 of consensus"  
 45798. .45909  
 repeat\_region /note="MTR1F repeat: matches 90. .192 of consensus"  
 46826. .46871  
 /note="23 copies 2 mer gt 97% conserved"  
 50684. .50969  
 repeat\_region /note="AluSg repeat: matches 1. .287 of consensus"  
 50960. .51291  
 /note="AluSx repeat: matches 1. .312 of consensus"  
 52222. .52519  
 repeat\_region /note="AluSx repeat: matches 1. .300 of consensus"  
 54065. .54260  
 /note="L1M4 repeat: matches 3865. .4055 of consensus"  
 54261. .54432  
 /note="FAM repeat: matches 2. .167 of consensus"  
 54433. .54629

```

/notes="LIM4 repeat: matches 3652..3865 of consensus"
repeat_region 54648..54862 repeat: matches 11..217 of consensus"
/notes="LIR4 repeat: matches 1..365 of consensus"
repeat_region 54863..55236
/notes="MTL1 repeat: matches 1..365 of consensus"
repeat_region 55237..55700

Query Match 67.9%; Score 808.8; DB 9; Length 16510;
Best Local Similarity 98.6%; Pred No. 1.5e-191;
Matches 816; Conservative 0; Mismatches 12; Indels 0; Gaps 0.

```

QY	254	CAGAGCTTCTTCAATATGAGGGCCATGGTGTGCTCCCCAGAGGGAGTCTGTGGCCCCCGCAGCAG	313
Db	90069	CAGAGCTTCTTCAACAGGGGCCATGAGTGTCTCCCAAGGGAGTCTGTGGCCCCCGCAGCAG	90128
QY	314	CAGCAGGTGCACAGCGTGGGTGTGTCTAGTCCCTCCCTCAATGACTCAACCGGACGTG	373
Db	90129	CAGCAGGTGCCAGCGTGGGTGTGTCTAGTCCCTCCCTCAATGACTCAACCGGACGTG	90188
QY	374	GTGAACATGGGCTTTCCGAGTGGCATCTCGGACACATGCTGTGGAGCGCGGTGACCTCC	433
Db	90189	GTGAACATGGGCTTTCCGAGTGGCATCTCGGACACATGCTGTGGAGCGCGGTGACCTCC	90248
QY	434	ATCTGTCTCTCTTCCGTGCATGAGTGTGGTGTCTGTGGCTCTCTCTGTGGTGTCTT	493
Db	90249	ATCTGTCTCTCTTCTCTGTCTCATGATGCTGTGGTGTCTGTGGCTCTCTCTGTGGTGTCTT	90308
QY	494	GTCTACCTGTGTGTCCACCTGAGTACCGGTGACTCTTGAAGGCTATAGGGGTGGTTT	553
Db	90309	GTCTACCTGTGTGTCCACCTGAGTACCGGTGACTCTTGAAGGCTATAGGGGTGGTTT	90368
QY	554	GTTGACAGGGACTTTGCTGGGCTTTGGTGTGAGACAGGCAATTGAGAGGGATCTGGTG	613
Db	90369	GTTGACAGGGACTTTGCTGGGCTTTGGTGTGAGACAGGCAATTGAGAGGGATCTGGTG	90428
QY	614	GTGCTTGAAGGTATGATCAGAGAGGGGACCAAGGTGTGTGTTCCTTGTGTGAAG	673
Db	90429	GTGCTTGAAGGTATGATCAGAGAGGGGACCAAGGTGTGTGTTCCTTGTGTGAAG	90488
QY	674	CGTAGGCGAGAGGAGACGTTAGTCCAGCATTTTCCAAAGTGTGGGTGGTCTGTGGTTC	733
Db	90489	CGTAGGCGAGAGGAGACGTTAGTCCAGCATTTTCCAAAGTGTGGGTGGTCTGTGGTTC	90548
QY	734	CCGAGATACCTTTTAGAGTGATGAGGGGCGCTCATTAAGTGCACAAATCAGACCAAGAA	793
Db	90549	CCGAGATACCTTTTAGAGTGATGAGGGGCGCTCATTAAGTGCACAAATCAGACCAAGAA	90608
QY	794	GCGATGCCCTTCCCAATCTCTCAATCCTTTTATGCCGAGAGATCTCAGCTGGATGCC	853
Db	90609	GCGATGCCCTTCCCAATCTCTCAATCCTTTTATGCCGAGAGATCTCAGCTGGATGCC	90668
QY	854	ACATGTTCCGATGCGCTGTGGAAGCATGCCGACGTCCTCCCTGATAGGGAGGAGACTT	913
Db	90669	ACATGTTCCGATGCGCTGTGGAAGCATGCCGACGTCCTCCCTGATAGGGAGGAGACTT	90728
QY	914	GGGCTTGGGGCAGGTGGAAGAAATTCAGACTTTTATAGCACTGTTTTGTTTTAAATGCT	973
Db	90729	GGGCTTGGGGCAGGTGGAAGAAATTCAGACTTTTATAGCACTGTTTTGTTTTAAATGCT	90788
QY	974	ATATTTTTATTGGCTACTTATTGTGTTAGCAAGTGTGTGTGCAATTCATTTATTATGTC	1033
Db	90789	ATATTTTTATTGGCTACTTATTGTGTTAGCAAGTGTGTGTGCAATTCATTTATTATGTC	90848
QY	1034	ACCTTTTCAATTAATAGTTTAACTTAAAGAAAAAATTTAAAAAATTTAAAAAATTTAAAAA	1081
Db	90849	ACCTTTTCAATTAATAGTTTAACTTAAAGAAAAAATTTAAAAAATTTAAAAAATTTAAAAA	90896

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

AC021954 GI:7417809  
AC021954.3 HTGS PHASE1; HTGS\_DRAFT.  
Homo sapiens (human)  
Homo sapiens  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Eumalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi;  
1 (bases 1 to 173341)  
Blair, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 10, clone RP11-57E12  
Unpublished

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases to 173341)  
 Birren, B., Lincoln, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,  
 Boguslavsky, L., Bokkigalter, B., Brown, A., Burnett, G., Castle, A.,  
 Coopers, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
 Dearbellano, K., Dewar, K., Domino, M., Doyle, M., Feneester, J.,  
 Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,  
 Gardyna, S., Grant, G., Hagsb, B., Heatard, A., Horton, A.,  
 Howland, J. C., Johnson, R., Jones, C., Kamm, L., Karatas, A., Klein, J.,  
 Landers, T., Lehocsky, J., Levine, R., Lien, C., Liu, G., Locke, K.,  
 McDonald, P., Marcus, N., McLean, P., McGuirk, A., McKernan, K.,  
 McPheters, R., Meldrum, J., Menais, L., Morrow, J., Naylor, J.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., Olivar, T. M., Peterson, K.,  
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thumann, N.,  
 Stojanovic, N., Subramanyam, A., Talamas, J., Tefaye, S., Theodore, J.,  
 Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, A., Wyman, D., Ye, W., J.  
 Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

**Direct Submission**  
Submitted (24-**Aug-2022**) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 5, 2000 this sequence version replaced gi:16984451.  
RefSeq accession number: [NC\\_000962.2](#)

Smt, A.T. & Green, E. 2000. <http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information

RESULT 5			
AC021954/c			
LOCUS			
AC021954	17341 bp	DNA	linear
			HTG 24-AUG-2002

```

----- Summary Statistics -----
Center ClonE name: 57_B_12
Sequencing vector: M13; M77815; 100% of reads
Chemistry Dye-terminator Big Dye; 100% of reads

```

NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1575	1574: contig of 1574 bp in length
*	1575	1674: gap of 100 bp
*	1675	3043: contig of 1369 bp in length
*	3044	3143: gap of 100 bp
*	3144	5889: contig of 2746 bp in length
*	5890	5899: gap of 100 bp
*	5900	8919: contig of 2990 bp in length
*	8980	9079: gap of 100 bp
*	9080	13674: contig of 4595 bp in length
*	13675	13774: gap of 100 bp
*	13775	18831: contig of 5057 bp in length
*	18832	18931: gap of 100 bp
*	18932	23562: contig of 4595 bp in length
*	23527	23562: gap of 100 bp
*	23527	27366: contig of 3760 bp in length
*	23627	27486: gap of 100 bp
*	27387	32572: contig of 5086 bp in length
*	27487	32672: gap of 100 bp
*	32573	38652: contig of 5960 bp in length
*	32673	38653: gap of 100 bp
*	38633	38752: gap of 100 bp
*	38733	43735: contig of 5003 bp in length
*	43736	44835: gap of 100 bp
*	43736	49020: contig of 5185 bp in length
*	43836	49120: gap of 100 bp
*	49021	53660: contig of 4540 bp in length
*	49121	53760: gap of 100 bp
*	53661	59544: contig of 5784 bp in length
*	53761	59545: gap of 100 bp
*	55545	59644: gap of 100 bp
*	55645	66335: contig of 6751 bp in length
*	66396	66495: gap of 100 bp
*	66496	74666: contig of 8201 bp in length
*	74697	74796: gap of 100 bp
*	74797	81336: contig of 6530 bp in length
*	81327	81426: gap of 100 bp
*	81427	89446: contig of 8020 bp in length
*	89447	89546: gap of 100 bp
*	89547	96459: contig of 6913 bp in length
*	96460	96559: gap of 100 bp
*	96560	106448: contig of 9939 bp in length
*	106499	106598: gap of 100 bp
*	106599	116966: contig of 10388 bp in length
*	116987	117086: gap of 100 bp
*	117087	128890: contig of 11804 bp in length
*	117088	128980: gap of 100 bp
*	128891	147300: contig of 18300 bp in length
*	128892	147300: gap of 100 bp
*	147291	147300: contig of 100 bp
*	147391	173341: contig of 25951 bp in length

```
misc_feature
source
Location/Qualifiers
1..173341
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/map="10"
/clone="RP11-57E12"
/clone_1fb="RP11-57E12"
1..1574 Human Male BAC
```

[illegible]

Db	160849	gcttacctggtgtccaccttgatgtaacgagatcctctgagggcctgattaggagtggttt	160799
QY	554	gttgagagagacattgcttgaggcccttggtgttgagagacagacatatattggaaggagatcttggtg	613
Db	160789	gttgagagagagacattgcttgaggcccttggtgttgagagacagacatatatttggaggagactctggtg	160730
QY	614	gttcctctttaaaggtatgatcagagaggggagccacagagtgctgtgtttccctctttgtgttaag	673
Db	160729	gttcctctttaaaggtatgatcagagaggggagccacagagtgctgtgtttccctctttgtgttaag	160670
QY	674	cgtgagagcagagggagacgttagtcacagatttccaaagtgtgggtgggtccgttggttc	733
Db	160669	cgtgagagcagagggagacgttagtcacagatttccaaagtgtgggtgggtccgttggttc	160610
QY	734	ccgagatpactttttaggtgtgtatggggcctgcacattaaattgagcacaagagagaaa	793
Db	160609	ccgagag-accttttaggtgtgtatggggcctgcacattaaattgagcacaagagagagaaa	160551
QY	794	ggcagtcgcccttcccaattctctcaattcctttatgscgagaaagctcagctgtagtcca	853
Db	160550	ggcagtcgcccttcccaattctctcaattcctttatgscgagaaagctcagctgtagtcca	160491
QY	854	acatgtttccgattccctgtgtgaaagacatggccagcgtctctctgcctaggagacagactt	913
Db	160490	acatgtttccgattccctgtgtgaaagacatggccagcgtctctctgcctaggagacagactt	160431
QY	914	ggagcttagggcagagtgagaaaaaattccacacttttttgacactggttttgattaatggt	973
Db	160430	ggagcttagggcagagtgagaaaaaattccacacttttttgacactggttttgattaatggt	160371
QY	974	atatttttattggcctactttatttggtttagacaaagtgtagtgacattctattattatgtg	1033
Db	160370	atatttttattggcctactttatttggtttagacaaagtgtagtgacattctattattatgtg	160311
QY	1034	accttttcantaaatgatatttaagtraaaaaaaaaaaaaaaaaaaaa	1081
Db	160310	accttttcantaaatgatatttaagtraaaaaaaaaatgagtcacattatmaagaaa	160263
RESULT 6			
LOCUS	AR413227	483 bp	DNA
DEFINITION	Sequence 864 from patent US 6639063.		linear
ACCESSION	AR413227		
VERSION	AR413227.1		
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
FEATURES	Unclassified.		
AUTHORS	1 (bases 1 to 483)		
TITLE	Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.		
JOURNAL	EST's and encoded human proteins		
FEATURES	Patent: US 6639063-A 864 28-Oct-2003;		
	Location/Qualifiers		
	1..483		
	/organism="unknown"		
	/mol_type="genomic DNA"		

Dp		131	TTCGTGTCACCCATCATGAGGACCACAGGGCTCGGAGATGTGCGAATTCTTGCCAA	190
Qy		181	TGGGAATTCGTGCAGACACCAACCCCGAGTGAAGAACCACTACCCAGCACCAAGG	240
Dp		191	TGGGGAATTCGTGCAGACACCAACCCCGAGTGAAGAACCACTACCCAGCACCAAGG	250
Qy		241	TAGCATTTCCTGACAAGAGCTCTTTCAAATAGSAGCCATGSGTCTCCCCAGGAGGTCTGG	300
Dp		251	TAGCATTTCCTGACAAGAGCTCTTTCAAATAGGAGGCCACTGATGTCTCCCCAGGAGGTCTGG	310
Qy		301	CCCCCGCAGACAGACGAGCGAGGTGCAGAGTGGTGTCTCACTGTCCTCCCTTCAATGACT	360
Dp		311	CCCCCGCAGACAGACGAGCGAGGTGCAGAGTGGTGTCTCACTGTCCTCCCTTCAATGACT	370
Qy		361	CACCGGAGAGTGTGTAACAATGGGCTTTCGGACGTGGACATCTCGGCAACATGCTGTGA	420
Dp		371	CACCGGAGAGTGTGTAACAATGGGCTTTCGGACGTGGACATCTCGGCAACATGCTGTGA	430
Qy		421	GCCGGTAGCCTCAATCTGTCTCTCTTCTGCTCATGATGCTTGGTTCGTG	473
Dp		431	GCCGGTAGCCTCAATCTGTCTCTCTTCTGCTCATGATGCTTGGTTCGTG	483
RESULT 7				
LOCUS	BD108780		483 bp	DNA linear PAT 18-SEP-2002
DEFINITION	EST and encoded human protein.			
ACCESSION	BD108780			
VERSION	BD108780.1 GI:23203598			
KEYWORDS	JP 2002010789-A/857. Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.			
AUTHORS	EST and encoded human protein			
TITLE	Patent: JP 2002010789-A 857 15-JAN-2002;			
JOURNAL	GENSET CORP			
COMMENT	OS Homo sapiens (human) PN JP 2002010789-A/857 PD 15-JAN-2002 PF 07-AUG-2000 JP 2000280898 PR 05-AUG-1999 US 60/1147499 PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT,JEAN EVE GIORDANO PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, C12N1/21, PC C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, C12N15/00 CC EST and encoded human protein FH Key Location/Qualifiers FT CDS Location/Qualifiers 1..483 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"			
FEATURES				
source				

[illegible]

	Query Match	39.7%	Score 473;	DB 6;	Length 483;
	Best Local Similarity	100.0%	Pred. No. 7,9e-108;		
	Matches 473;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AAACCGCCGAGAGCCGCCGAGGTGACGCGCCGAGAGCCCGGAGTGGCGCTGCAGACA	60		
Db	11	AAACCCCTCGAGAGCCGCCGAGGTGACGCGCCGAGAGCCCGGAGTGGCGCTGCAGACA	70		
Qy	61	CAGCATCTACTAGCGTGGGTCACTCTGTGAAATATCACTGACTGCAAGCCTTCCTCAAT	120		
Db	71	CAGCATCTTACTAGGTGGGTCACTCTGGAACATATCACTGACTGCAAGCCTTCCTCAAT	130		

QY 121 TTCTGTGAGCCCATCAGGSAACCAAGCGCTGGGAGATGTGCGANTCTTGGCCA 180  
Db 131 TTCTGTGAGCCCATCAGGSAACCAAGCGCTGGGAGATGTGCGANTCTTGGCCA 190  
QY 181 TGGGAAATCTGTGAGAGAGAGACCCCGAGTGAAGACCACTACCAAGCCACAGAG 240  
Db 191 TGGGAAATCTGTGAGAGAGAGACCCCGAGTGAAGACCACTACCAAGCCACAGAG 250  
QY 241 TAGCATTTCTCGACAGAGCTTCTTAATAGGGGACATGGTGTCCCGAGGGGCTCTGG 300  
Db 251 TAGCATTTCTCGACAGAGCTTCTTAATAGGGGACATGGTGTCCCGAGGGGCTCTGG 310  
QY 301 CCCCCCAGAGAGAGAGAGTCCAGGCTGGTGTCTCAATGCCCTTCAATGACT 360  
Db 311 CCCCCCAGAGAGAGAGTCCAGGCTGGTGTCTCAATGCCCTTCAATGACT 370  
QY 361 CAACCGGAGCTGTGTAACATGGGCTTTCGAGTGGACATCGGCAACATGCTGTGA 420  
Db 371 CAACCGGAGCTGTGTAACATGGGCTTTCGAGTGGACATCGGCAACATGCTGTGA 430  
QY 421 GCCGCTGACCTCCATCTGCTCTTCTCTGCTCATGATGCTGTGTGTG 473  
Db 431 GCCGCTGACCTCCATCTGCTCTTCTCTGCTCATGATGCTGTGTGTG 483

RESULT 8  
AX11082 550 bp DNA linear PAT 01-MAY-2001  
LOCUS Sequence 5 from Patent WO0127276.  
DEFINITION AX11082  
ACCESSION AX11082  
VERSION AX11082.1 GI:13939514  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 Harlocker, S.L., Dillon, D.C. and Xu, J.  
DNA sequences from breast tumor and uses thereof  
Patent: WO 0127276-A 5 19-APR-2001;  
CORIXA CORPORATION (US)  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..550  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..550  
/note="n = A,T,C or G"

ORIGIN  
misc\_feature  
Query Match 38.0%; Score 452.6; DB 6; Length 550;  
Best Local Similarity 95.4%; Pred. No. 1.1e-102; Indels 8; Gaps 5;  
Matches 519; Conservative 0; Mismatches 17;

QY 47 CGCGCTGACAGACAGATCTACTAGCGGTGGGTCACTGTGAAATCACTGACTGC 106  
Db 1 CGCGCTGACAGACAGATCTACTAGCGGTGGGTCACTGTGAAATCACTGACTGC 60  
QY 107 AAGCTTCCTCAATTTCTGTGTGAGAGCCCATCAGGACCCACAGCGCTTGGAGATGTG 166  
Db 61 AAGCTTCCTCAATTTCTGTGTGAGAGCCCATCAGGACCCACAGCGCTTGGAGATGTG 120  
QY 167 CGGATCTTGGCCATATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 226  
Db 121 CGGATCTTGGCCATATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
QY 227 CAGCCACCAAGAGATGATCTTCTGACAGAGCTTCTTAATAGGGGACATGGTGTCTCC 286  
Db 181 CAGCCACCAAGAGATGATCTTCTGACAGAGCTTCTTAATAGGGGACATGGTGTCTCC 240  
QY 287 CAGAGGGGCTCTGGCCCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 346  
Db 241 CAGAGGGGCTCTGGCCCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

QY 347 CCCTTCAATGACTCAACCGGACAGTGTGAACATGGGCTTTCCGAGATGATCTCGGC 406  
Db 301 CCCTTCAATGACTCAACCGGACAGTGTGAACATGGGCTTTCCGAGATGATCTCGGC 360  
QY 407 AACCATGC-TGTGAGCCCGGTGACCTCATCTGCTCTTCTCTGCTCATGATGCTTGG 465  
Db 361 AACCATGCTGTGAGCCCGGTGACCTCATCTGCTCTTCTCTGCTCATGATGCTTGG 419  
QY 466 TGTGTGAGGCTCTCCCTGTTGGCTGTGTACTAGTGTGTGTCACAGCTGATGTC-AGGGGT 524  
Db 420 TGTGTGAGGCTCTCCCTGTTGGCTGTGTACTAGTGTGTGTCACAGCTGATGTCAGGGGT 479  
QY 525 GACCTCT--GAGGAGCTGATAGG--GTGGCTTGTGTGAGAGGAGCTTGTGAGCTTGG 579  
Db 480 GACCTTTTGAAGGAGCTGATAGGAGGTGGGCTTGTGTGAAAGAACTTGTGAGCTTGG 539  
QY 580 TGTG 583  
Db 540 TGGG 543

RESULT 9  
AX431087 643 bp DNA linear PAT 28-JUN-2002  
LOCUS Sequence 30 from Patent WO0240535.  
DEFINITION AX431087  
ACCESSION AX431087  
VERSION AX431087.1 GI:21656064  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 Resipon, H., Liu, C., Salceda, S., Sun, Y., Caffery, R. and  
Macina, R.A.  
Compositions and methods relating to ovary specific genes and  
proteins  
Patent: WO 0240535-A 30 23-MAY-2002;  
DIADEXUS INC (US)  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..643  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 36.3%; Score 432.8; DB 6; Length 643;  
Best Local Similarity 96.1%; Pred. No. 1.1e-97; Indels 9; Gaps 7;  
Matches 519; Conservative 0; Mismatches 12;

QY 662 CTTTGTGTTAAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
Db 104 CTTTGTGTTAAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 163  
QY 721 GGTCCGTTGTTCCCGAGATCTTTTAGTGTGTATGGGCTTGCATTAAGTGGAC-AAA 779  
Db 164 GGTCCGTTGTTCCCGAGATCTTTTAGTGTGTATGGGCTTGCATTAAGTGGACAAA 223  
QY 780 ATCAGAGCAAGAAAGCGATGCTTT-CCCATTTCTTCAATCTTTT--ATGCCGAGAG 836  
Db 224 ATCAGAGCAAGAAAGCGATGCTTTCCCATTTCTTCAATCTTTTATGGCCGAGAG 283  
QY 837 ATTCAGCTGATGCAACATGTTCCAGTGCCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 896  
Db 284 ATTCAGCTGATGCAACATGTTCCAGTGCCTGTGAGAGAGAGAGAGAGAGAGAGAGAG 343  
QY 897 CCTAG 954  
Db 344 CCTAG 403  
QY 955 CTGTTTTTCTTATAGTATATT-TTATTTGGTACTTTATTTGTTTGAAGAGAGTGTGA 1013

Db 404 TGTATTTTGTATATGATATTTCTATATGCTACTTTATTTATAGACAAGTGTA 463  
QY 1014 GTGGCATTTCTATTTATT-GTACCTTTTCAATAATAGATTTTAAGTAAAAA 1072  
Db 464 GTGGCATTTCTATTTATTGTGACCTTTTCAATAATAGATTTTAAGCAAAAA 523  
QY 1073 AAAAAAAAAAAGCTGTGGGGGTTTACCGGGCCAAAGGGGGCCCCGGGGGGAATCT 1132  
Db 524 AAAAAAAAAAAGCTGTGGGGGTTTACCGGGCCAAAGGGGGCCCCGGGGGGAATCT 583  
QY 1133 GGTTCCTCCGCCCCCAATTCCGCCCAATTTTGTACAAATGAAGAGACACAGAAACC 1192  
Db 584 GGTTCCTCCGCCCCCAATTCCGCCCAATTTTGTACAAATGAAGAGACACAGAAACC 643

RESULT 10  
AR415427 519 bp DNA linear PAT 18-DEC-2003  
LOCUS AR415427  
DEFINITION Sequence 3064 from patent US 6639063.  
ACCESSION AR415427  
VERSION AR415427.1 GI:40170537  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1. (bases 1 to 519)  
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.  
TITL E EST's and encoded human proteins  
JOURNAL Patent: US 6639063-A 3064 28-OCT-2003;  
FEATURES  
source 1..519  
location/Qualifiers  
1..519  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 32.9%; Score 391.8; DB 6; Length 519;  
Best Local Similarity 98.8%; Pred. No. 2.1e-87;  
Matches 404; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
QY 54 GCAGACACAGCATTTCTACTGAGCGTGGGTCACTCTGTGAAATCACTGACTGCAAGCTTC 113  
Db 112 GAATACACACATTTCTACTGAGCGTGGGTCACTCTGTGAAATCACTGACTGCAAGCTTC 171  
QY 114 CTTCAATTTCTGTGTGACGCCCATCAGGAGCCCAAGCGCTGGAGGATGTGTGGATCT 173  
Db 172 CTTCAATTTCTGTGTGACGCCCATCAGGAGCCCAAGCGCTGGAGGATGTGTGGATCT 231  
QY 174 TGGCCAAATGGGAAATGTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 233  
Db 232 TGGCCAAATGGGAAATGTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 291  
QY 234 CAAGAGGTAGCATTTCTCTGACAGAGCTTTCTTAATAGGGCCATGTGTCTCCCAAGGG 293  
Db 292 CAAGAGGTAGCATTTCTCTGACAGAGCTTTCTTAATAGGGCCATGTGTCTCCCAAGGG 350  
QY 294 GTCTGTGGCCCCCGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 353  
Db 351 GTCTGTGGCCCCCGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 410  
QY 354 ATGACCTCAACCGGACAGCTGTGAACATGGGCTTTCCGAGTGGCATCTCGCAACCATG 413  
Db 411 ATGACCTCAACCGGACAGCTGTGAACATGGGCTTTCCGAGTGGCATCTCGCAACCATG 470  
QY 414 CTGTGAGCGGTGACCTTCATCTGTCTCTTTCTCTGCTCATGATGCT 462  
Db 471 CTGTGAGCGGTGACCTTCATCTGTCTCTTTCTCTGCTCATGATGCT 519

RESULT 11  
BD110980 519 bp DNA linear PAT 18-SEP-2002  
LOCUS BD110980  
DEFINITION EST and encoded human protein.

VERSION BD110980.1 GI:23205798  
KEYWORDS JP 2002010789-A/3057.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1. (bases 1 to 519)  
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.  
TITL E EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 3057 15-JAN-2002;  
GENSET CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002010789-A/3057  
PD 15-JAN-2002  
PF 07-AUG-2000 JP 2000280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12Q1/68,C12N1/15,C12N1/19,PC  
C12N1/21,  
PC C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N1/15,C12N1/19,PC  
C12N15/00  
CC EST and encoded human protein  
FH Key Location/Qualifiers  
FT CDS 93..431.  
Location/Qualifiers  
1..519  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 32.9%; Score 391.8; DB 6; Length 519;  
Best Local Similarity 98.8%; Pred. No. 2.1e-87;  
Matches 404; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
QY 54 GCAGACACAGCATTTCTACTGAGCGTGGGTCACTCTGTGAAATCACTGACTGCAAGCTTC 113  
Db 112 GAATACACACATTTCTACTGAGCGTGGGTCACTCTGTGAAATCACTGACTGCAAGCTTC 171  
QY 114 CTTCAATTTCTGTGTGACGCCCATCAGGAGCCCAAGCGCTGGAGGATGTGTGGATCT 173  
Db 172 CTTCAATTTCTGTGTGACGCCCATCAGGAGCCCAAGCGCTGGAGGATGTGTGGATCT 231  
QY 174 TGGCCAAATGGGAAATGTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 233  
Db 232 TGGCCAAATGGGAAATGTGTGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 291  
QY 234 CAAGAGGTAGCATTTCTCTGACAGAGCTTTCTTAATAGGGCCATGTGTCTCCCAAGGG 293  
Db 292 CAAGAGGTAGCATTTCTCTGACAGAGCTTTCTTAATAGGGCCATGTGTCTCCCAAGGG 350  
QY 294 GTCTGTGGCCCCCGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 353  
Db 351 GTCTGTGGCCCCCGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 410  
QY 354 ATGACCTCAACCGGACAGCTGTGAACATGGGCTTTCCGAGTGGCATCTCGCAACCATG 413  
Db 411 ATGACCTCAACCGGACAGCTGTGAACATGGGCTTTCCGAGTGGCATCTCGCAACCATG 470  
QY 414 CTGTGAGCGGTGACCTTCATCTGTCTCTTTCTCTGCTCATGATGCT 462  
Db 471 CTGTGAGCGGTGACCTTCATCTGTCTCTTTCTCTGCTCATGATGCT 519

RESULT 12  
G22561 480 bp DNA linear STS 31-MAY-1996  
LOCUS G22561  
DEFINITION human STS WI-12709, sequence tagged site.  
ACCESSION G22561  
VERSION G22561.1 GI:1342887  
KEYWORDS STS; STS sequence; primer; sequence tagged site.



Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLAST at: <http://image.llnl.gov>  
Series: IRAC Plate: 36 Row: h Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.  
Location/Qualifiers

FEATURES  
source

1. 1538  
/organism="Mus musculus"  
/mol\_type="cDNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="MGC:28201 IMAGE:3989515"  
/tissue\_type="mammary tumor metastasized to lung, tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."  
/clone\_11b="MGI CGAP\_1u29"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1. 1538  
/gene="pCMV-SPORT6"  
/db\_xref="locusid:69894"  
/db\_xref="mgi:1917144"  
486..848  
/codon\_start=1  
/product="pCMV-SPORT6 protein"  
/protein\_id="AAH24943.1"  
/db\_xref="GI:19354289"  
/db\_xref="locusid:69894"  
/translation="MVRILNGEIVDDDDPRVTTTTHSSSSQGFNRRGAPRG  
PGRRODAGRLAAGSPFSDLRQLVNNFPMHGNVBEVTISLLFLMLGV  
RGLLVLTVALVSHLSOR"

## ORIGIN

Query Match 25.8%; Score 307.4; DB 10; Length 1538;  
Best Local Similarity 72.6%; Pred. No. 4,7e-66;  
Matches 484; Conservative 0; Mismatches 166; Indels 17; Gaps 6;

119 ATTTCTGTGACGCTATCAGGAGCCAGCG-CTGGAGAGTGGTGGATCTTGGC 177  
443 ATTACAGATCTGACAGCAGAAATCCAGCGTCCAGGAGAAATGATGATCTTGGC 502  
178 CAATGGGGAATCGTGCAGGAGCGAGCCCGAGTGGAGACACATCCAGGAGCAAG 237  
503 CAATGGGGAATCGTGCAGGAGCGAGCCCGAGTGGAGACACATCCAGGAGCAAG 559  
238 AGTAGATCTCTGACAGAGCTTTCAATAGGAGCGAGTGGTCTCCCGAGGAGTCC 297  
560 AAGTAGAGCTCTCAGAGAGGCTTTTCAACAGAGCGCAGCGGCACTCCAGGAGTCC 619  
298 TGGCCCCCGCAGAGAGGAGCGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 357  
620 TGAACCCCGCAGAGAGGAGCGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 679  
358 CCTCAACCCGAGTGGTGAATAGGAGCTTCCGAGTGGATCTCCGCAACCATCTGT 417  
680 CCTCAACCCGAGTGGTGAATAGGAGCTTCCGAGTGGATCTCCGCAACCATCTGT 739  
418 GGAACCCGAGTGGTGAATAGGAGCTTCCGAGTGGATCTCCGCAACCATCTGT 477  
740 GGAACCCGAGTGGTGAATAGGAGCTTCCGAGTGGATCTCCGCAACCATCTGT 799  
478 CCTCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 536  
800 CCTCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 859  
537 CTGATAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 591  
860 CTGATAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 919  
592 CATTTTGAAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 651

Db 920 CACAGGAGGAGGAGCTTGGTGGGCTCAGAGGTGTGATGGAGTACAGCCGGCTTGT 979  
QY 652 TGTGTTTCCCTTTGTGTTAGCTGAGCGCAGAGGAGACCTTATGTCAGATTTCCAAA 711  
Db 980 AGATTTTCTCCCGT-----GCAGAGGAGGAGGAGCTCTCTGGCCAGATTTGCGAA 1034  
QY 712 GTGGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 771  
Db 1035 --AAAGAGAGTCTGTGTTCTGTGATGTTTATGAGGAGGAGTGGGCTTGGCTTAAA 1092  
QY 772 GGCACAA 778  
Db 1093 AGCACA 1099

RESULT 14  
AC116233/c  
LOCUS  
DEFINITION  
Rattus norvegicus clone CH230-95L17, \*\*\* SEQUENCING IN PROGRESS  
AC116233  
275924 bp DNA linear HTG 22-SEP-2002  
Rattus norvegicus  
AC116233.7 GI:23101097  
HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_ENRICHED.  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 275924)  
Munzy, D., Marie, M., Metzger, M., Lee, A., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alebrooks, S., Amin, A., Anguiano, D.,  
Aryaloechi, V., Ayagi, A., Ayodeji, M., Bacc, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benamed, F.,  
Bismail, K., Blair, V., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, Z.,  
Chacko, J., Chavez, D., Chen, R., Chen, R., Chen, Y., Chen, Z.,  
Cleveland, L., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Devila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Deigado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C., Fails, T., Fan, G.,  
Fernandez, S., Finley, A., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Gatta, M.,  
Gebregorjse, E., Geer, K., Gill, R., Gilly, M., Guerra, M., Guervara, W.,  
Gumarane, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S., Hodgson, A., Hogues, M.,  
Hollins, B., Howells, S., Huix, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowals, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorensuhera, L., Louissege, H., Lorado, R., Lu, X., Ma, J.,  
Maneswar, M., Mahindaratne, M., Mamoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhinley, S., McLeod, M., McNeill, T., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, N., Morris, K., Morris, S., Muniz, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwokelemeh, O., Okonou, G., Olariunsaogon, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Plankoch, C.,  
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,  
Puzio, B., Quiroz, J., Rachin, E., Reeves, K., Regier, M., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,  
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., Smaj, D.,  
Sneid, A., Sodergren, E., Song, X., Sotelle, R., Sosa, J.,  
Stedme, M., Strong, R., Sutton, A., Sytek, A., Taber, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmami, K.,  
Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, D.,





LOCUS AC111675 240657 bp DNA 1linear HTG 13-MAY-2003  
 DEFINITION Rattus norvegicus clone CH230-48G14, WORKING DRAFT SEQUENCE, 4  
 UNORDERED PIECES.  
 AC111675  
 HTG: HTG PHASE: HTGS DRAFT; HTGS\_FULLTOP.  
 Rattus norvegicus (Norway rat)  
 Rattus norvegicus  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 240657)  
 Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alshrooke, S., Amin, A., Anguiano, D.,  
 Ayala-Bech, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaru, D., Barker, M., Barnstead, M., Benahmed, F.,  
 Bismuth, K., Blair, J., Blankenship, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M., Davis, C., Davy, Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,  
 Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Fraser, C., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gheorghe, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, J.,  
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladik, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, D., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kwis, C., Kraft, C., Lebowitz, H., Levan, L., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorenz, W., Louie, S., Lozada, R., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
 Mawhinney, S., McLeod, M., McNeill, T., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Muidasa, M., Murphy, M., Nair, L.,  
 Narkiewicz, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nowakowski, O., Okwum, G., Olariu, S., Pal, S., Parks, K.,  
 Patel, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Ploger, F., Poldoski, A., Popovic, D., Primus, E., Pu, L.,  
 Puzo, M., Quirio, J., Rachin, E., Reeves, K., Regier, M., Reigh, R.,  
 Rell, B., Rell, M., Ren, Y., Reuter, R., Richards, S., Riggs, F.,  
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,  
 Sanders, M., Savary, G., Scherer, S., Scott, C., Shlitsman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., Smaj, D.,  
 Sneed, A., Sodergren, E., Song, X., Sorelle, R., Sosa, J.,  
 Steinle, M., Strong, R., Sutton, A., Svalbe, A., Taber, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
 Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Wilson, R., Wlezyk, R., Wooden, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhu, D., von  
 Niederhausen, A., Weiss, R., Smith, D., Holt, R., Smith, H., O.,  
 Weinstock, G., and Gibbs, R. A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 240657)  
 Worley, K. C.  
 Submitted (19-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 240657)  
 Rat Genome Sequencing Consortium.  
 Direct Submission

## COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On May 13, 2003 this sequence version replaced gi:2481892.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.

## Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GMV1

Center clone name: CH230-48G14

Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 228778 bases at least Q40

Consensus quality: 231963 bases at least Q30

Consensus quality: 234139 bases at least Q20

Estimated insert size: 23755; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length

\* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 4 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 236811: contig of 236811 bp in length

\* 236812 236911: gap of unknown length

\* 236912 238255: contig of 1344 bp in length

\* 238256 238355: gap of unknown length

\* 238356 239605: contig of 1150 bp in length

\* 239606 239607: gap of unknown length

\* 239608 240657: contig of 1052 bp in length.

## FEATURES

## source

/organism="Rattus norvegicus"

/mol\_type="Genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-48G14"

1. 2145

/note="wgs end extension

clone\_end:Sp6"

2196. 3484

/note="wgs end extension

clone\_end:Sp6"

8292. 9121

/note="clone boundary

clone\_end:Sp6

site:ECORI

end sequence: BH363877"

complement(233966..234793)

/note="clone boundary

clone\_end:17

site:ECORI

end\_sequence: BH363875"

234903..236811

/note="wgs end extension

clone\_end:T7"

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

ORIGIN

Query Match 24.7%; Score 294; DB 2; Length 240657;  
 Best Local Similarity 71.6%; Pred. No. 4.4e-62;  
 Matches 499; Conservative 0; Mismatches 181; Indels 17; Gaps 8;

QY 123 CTGGTGAAGCCATCAAGGACCCACAGGCTT-GGAGGATGATGCGATCTTGCCAAAT 181  
 DB 51661 CAGGTCCAGCCAGCAGAAATTCACAGGCTCTAGGATATGATGAGATCTTGCTAAAT 51602  
 QY 182 GGGGAAATGATGAGACGACGACCCCGAGTGAAGACCACTACCCAGCCACCAAGGT 241  
 DB 51601 GGGGAAATGATGAGATGATGACCCCGAGTGAAGACCACTACCCAGGAGT 51545  
 QY 242 AGCATTCCTGACAGAGCTTCTCAATAGGGGCCATGATGCTCCCGAGGGGCTCT-GG 300  
 DB 51544 AGCAGACCCCGAGGGGCTTTTCAACAGAGCCAAGTGCACCTCCAGGGGGCTTGAG 51485  
 QY 301 CCCCCCAGACAGCAGGACAGGCTGAGGCTGCTCACTCCCTTCAATGACTT 360  
 DB 51484 ACCCCAGCAGCAGCAGGACAGGCTGAGGCTGCTCACTCCCTTCAATGACTT 51425  
 QY 361 CAACCGGAGCTGATGAACATGAGGCTTTCGAGTGGACATCGGCAACCATGCTGGA 420  
 DB 51424 GAAACCGGAGCTGATGAACATGAGGCTTTCGAGTGGACATCTTGGAACCATGCTGGA 51365  
 QY 421 GCCGGTGAACCTCATCTGCTCTCTTCTGCTCATGATGCTGATGCTGCTGCT 480  
 DB 51364 GCCGTGACCTCATCTGCTCTCTTCTGCTCATGATGCTGAGGCTGCTGCTCT 51305  
 QY 481 CTTGTTGAGCTTGTCTACCTGATGCCATCCAGTCAAGGCTGACCTGAGGGCTGA 540  
 DB 51304 GCTTTGAGCTTGTCTACCTGATGCCATCCAGTCAAGGCTGACCTGAGGGCTGA 51245  
 QY 541 TAGGGGTGAGTTTGTGAGAGGACCTGCTGAGGCTTGTGATGAGAGCAGGATATTGG 600  
 DB 51244 TAGGCAGGACGATGTTCAAGTATGATGATGAGGCTTGTGATGAGAGCAGGATATTGG 51186  
 QY 601 AGGGGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 DB 51185 AGGGGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 51129  
 QY 661 CTTTGTGTTAAGCTGAGCAGAGGAGACGTTAGTCCAGCATTTCCAAAGTGTGGTG 720  
 DB 51128 CC-----CGCTGACAGAGGACGAGGACCTCTGAGGCGAGCATTTCTGAAAAATAG-A 51076  
 QY 721 GGTCCGTGTTCCCGAGATCTTTTAAAGTGTATGAGGAGCTGATTAAGTGGCAAAA 780  
 DB 51075 GTCCCATGTTTTCTGAAATGTTTTTAAAGGAGCATAGGCTCTCCATTACAAGCAC-AAAG 51017  
 QY 781 TCAGAGCAGAAAGCAGATGCCCTTCCCAATTCCTCA 817  
 DB 51016 CAGAGTAAAGAGGCTTAGTATGTTGATTCCTCA 50980

Search completed: August 22, 2004, 02:37:28  
 Job time : 4696 secs

**This Page Blank (uspto)**

Run on: August 21, 2004, 16:20:58 ; Search time 514 Seconds

9851.852 Million cell updates/sec

Title: US-10-001-885-31  
Perfect score: 1192

Sequence: 1 aaaccgctggagccgcgg...ggaagaggaacacgaaccc 1192

Scoring table:

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

```
Database : N_Geneseq_29Jan04:*
```

1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002s: \*  
7: geneseqn2003as: \*  
8: geneseqn2003bs: \*  
9: geneseqn2003cs: \*  
10: geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1192	100.0	1152	6	ABN87820	ABn87820 Human ova
2	1061.4	89.0	1108	6	ABL90057	ABl90057 Human pol
3	1012	84.9	1049	6	AAV95558	AAv95558 Human sec
4	1012	84.9	1049	6	ABSV3545	ABsv3545 Human CDN
5	1012	84.9	1049	8	ACD82688	ACd82688 CDNA sequ
6	562	47.1	562	6	ABZ11212	ABz11212 Human pol
7	455.4	38.2	1165	5	AAAS4820	AAe4820 DNA encod
8	452.6	38.0	550	5	AAE82505	AAe82505 Human bre
9	434.6	36.5	587	5	AAE64821	AAe64821 DNA encod
10	432.8	36.3	643	6	ABN87819	ABn87819 Human ova
11	372.2	31.2	517	7	ABBX4673	ABx4673 Human CDN
12	363.4	30.5	365	7	ABX92251	ABx92251 Human ova
13	363.2	30.5	612	2	ABX92250	ABx92250 Human ova
14	159	13.3	348	5	AAE64819	AAe64819 DNA encod
15	75.6	6.3	461	6	ABT07555	ABt07555 Human bre
16	75.6	6.3	633	6	ABT07560	ABt07560 Human bre
17	75	6.3	986	7	ABT13409	ABt13409 Breast sp
18	71.2	6.0	420	6	ABT07648	ABt07648 Human bre
19	71.2	6.0	2269	6	ABT07649	ABt07649 Human bre
20	69.6	5.8	577	6	ABX92021	ABx92021 Lung spec
21	67.8	5.7	355	7	ABX93192	ABx93192 Human ova
22	67.4	5.7	999	6	ABBS67265	ABbs67265 Breast sp
23	67	5.6	668	6	ABT04062	ABt04062 Human ova

## ALIGNMENTS

24	65.2	5.5	737	7	ABH13319	Breast sp
25	64.9	5.4	766	6	ABH07679	Human bre
26	64.4	5.4	1400	6	ABH07680	Human bre
27	64	5.4	666	6	ABH87812	Human ova
28	63.6	5.3	536	6	ABH07561	Human bre
29	63.6	5.3	626	6	ABH87873	Human ova
30	63.6	5.3	657	7	ABX92246	Human ova
31	63.6	5.3	1206	6	ABH07562	Human bre
32	62.8	5.3	626	6	ABH67291	Breast sp
33	62.2	5.2	716	6	ABH04069	Human ova
34	62.2	5.2	819	7	ABH13356	Breast sp
35	61.8	5.2	601	6	ABH83649	Human bre
36	61	5.1	422	4	AA115924	Human bre
37	60.4	5.1	800	7	ABX92232	Human ova
38	60.4	5.1	1545	6	ABX91992	Lung spec
39	59.8	5.0	1623	6	ADJ39107	Human lun
40	59.6	5.0	464	5	ABH093810	Human pro
41	59.2	5.0	816	6	ABH67266	Breast sp
42	58.4	4.9	810	6	ABH87806	Human ova
43	57.8	4.8	1455	7	ABX92216	Human ova
44	57.8	4.8	2910	6	ABH87829	Human ova
45	57.4	4.8	665	7	ABX92249	Human ova

CC	RESULT 1
ID	ABN87820
XX	ABN87820 standard; cDNA; 1192 BP.
AC	ABN87820;
DT	12-AUG-2002 (first entry)
XX	
DE	Human ovary specific nucleic acid SEQ ID NO:31.
XX	
KW	Human; ovary specific nucleic acid; OSNA; ovary specific protein; OSP;
KW	ovary specific gene; OSG; ovarian cancer; immune response; metastasis;
KW	chromosome 10; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200240535-A2.
XX	
PD	23-MAY-2002.
XX	
PP	20-NOV-2001; 2001WO-US045011.
XX	
FR	20-NOV-2000; 2000US-0252061P.
FR	27-NOV-2000; 2000US-0253257P.
XX	
PA	(DIAD-) DIADEXUS INC.
XX	
PI	Salceda S, Macina RA, Recipon H, Cafferey R, Sun Y, Liu C;
XX	
DR	WPI; 2002-471617/50.
XX	
PT	New ovary specific genes and proteins, useful as a vaccine for treating
PT	patients with ovarian cancer, or for diagnosing and monitoring the
PT	presence and metastases of ovarian cancer in a patient.
XX	
PS	Claim 1; Page 173-174; 260bp; English.
XX	
CC	ABN87790 to ABN87882 represent human ovary specific nucleic acid (OSNA)
CC	sequences, and ABN79297 to ABN79370 represent ovary specific protein
CC	(OSP) sequences from the present invention. OSNA and OSP sequences have
CC	cytotoxic activity, and can be used in vaccine production and gene
CC	therapy. An antibody that specifically binds to an OSP can be used for
CC	treating a patient with ovarian cancer, particularly for inducing an
CC	immune response against the ovarian cancer cell expressing the OSNA or
CC	OSP. The OSNAs and OSPs can also be used for diagnosing and monitoring
CC	the presence and metastases of ovarian cancer in a patient

XX	Sequence	1192 BP	262 A	298 C	349 G	283 T	0 U	0 Other
SQ								

Query Match	100.0%;	Score 1192;	DB 6;	Length 1192;
-------------	---------	-------------	-------	--------------

Best Local Similarity 100.0%; Pred. No. 5.4e-264;  
Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AAACCCGCTGAGAGCCGCGAGGTGAGCGCCCGAGGCTCGCGCTCGACACA	60
Db	1	AAACCCGCTGAGAGCCGCGAGGTGAGCGCCCGAGGCTCGCGCTCGACACA	60
QY	61	CACACATCTACTCAGCGGTGGGTCACTCTGTGACATCATCTGACTGCAAGCTTCCCTCAAT	120
Db	61	CACACATCTACTCAGCGGTGGGTCACTCTGTGACATCATCTGACTGCAAGCTTCCCTCAAT	120

QY	121	TTCTGCTGACGCCATTCAGGAGCCACAGCCCTCTGGAGGATGATGCGAGATCTTTGGCCAA	180
Db	121	TTCTGCTGACGCCCATCAGGAGCCACAGCCCTCTGGAGGATGATGCGAGATCTTTGGCCAA	180
QY	181	TGGGGAATCGTGCAGGAGCAGACGACCCCGAGTAGGAGCCATACCCAGCCACCAAGAG	240
Db	181	TGGGGAATCGTGCAGGAGCAGACGACCCCGAGTAGGAGCCATACCCAGCCACCAAGAG	240

QY	241	TAGCATTCTTCAGCAGAGCTTCTTCAATAGGGGCGCATGGTCTCCCCCAGGGGGTCTGG	300
Db	241 <td>TAGCATTCTTCAGCAGAGCTTCTTCAATAGGGGCGCATGGTCTCCCCCAGGGGGTCTGG <td>300</td> </td>	TAGCATTCTTCAGCAGAGCTTCTTCAATAGGGGCGCATGGTCTCCCCCAGGGGGTCTGG <td>300</td>	300
QY	301 <td>CCCCCCCCCAGCAGCAGCGCAGGTGCGAGTGGTGTCTCACTCCCTCCCTTAATGACT <td>360</td> </td>	CCCCCCCCCAGCAGCAGCGCAGGTGCGAGTGGTGTCTCACTCCCTCCCTTAATGACT <td>360</td>	360
Db	301 <td>CCCCCCCCCAGCAGCAGCGCAGGTGCGAGTGGTGTCTCACTCCCTCCCTTAATGACT <td>360</td> </td>	CCCCCCCCCAGCAGCAGCGCAGGTGCGAGTGGTGTCTCACTCCCTCCCTTAATGACT <td>360</td>	360

Qy	361	CAACCGGAGCTGTGTGAACATGAGCTTTCGCGAATGCGAACCATGCTGTGA	420
Db	361	CAACCGGAGCTGTGTGAACATGAGCTTTCGCGAATGCGAACCATGCTGTGA	420
Qy	421	GCGGAGACCTCATCCGCTGCTCTTCTCTCATATCTTGAGTTCGAGGCTCCT	480
Db	421	GCGGAGACCTCATCCGCTGCTCTTCTCTCATATCTTGAGTTCGAGGCTCCT	480

QY	481	CTGGATTGSCCTTGTCTACCTGGTGTGCCACTGAGTCAACGGGTGAACCTTGAGGGCTGA	540
Db <td>481 <td>CTGGATTGSCCTTGTCTACCTGGTGTGCCACTGAGTCAACGGGTGAACCTTGAGGGCTGA <td>540</td> </td></td>	481 <td>CTGGATTGSCCTTGTCTACCTGGTGTGCCACTGAGTCAACGGGTGAACCTTGAGGGCTGA <td>540</td> </td>	CTGGATTGSCCTTGTCTACCTGGTGTGCCACTGAGTCAACGGGTGAACCTTGAGGGCTGA <td>540</td>	540
QY <td>541 <td>TAGGGGTGGGTTTGTGTGAGAGGAACTTGTGGGCGCTTGTTGTGAGAGCAAGCATTTTGG <td>600</td> </td></td>	541 <td>TAGGGGTGGGTTTGTGTGAGAGGAACTTGTGGGCGCTTGTTGTGAGAGCAAGCATTTTGG <td>600</td> </td>	TAGGGGTGGGTTTGTGTGAGAGGAACTTGTGGGCGCTTGTTGTGAGAGCAAGCATTTTGG <td>600</td>	600
Db <td>541 <td>TAGGGGTGGGTTTGTGTGAGAGGAACTTGTGGGCGCTTGTTGTGAGAGCAAGCATTTTGG <td>600</td> </td></td>	541 <td>TAGGGGTGGGTTTGTGTGAGAGGAACTTGTGGGCGCTTGTTGTGAGAGCAAGCATTTTGG <td>600</td> </td>	TAGGGGTGGGTTTGTGTGAGAGGAACTTGTGGGCGCTTGTTGTGAGAGCAAGCATTTTGG <td>600</td>	600

QY 601 AGGGAGTCTGTGTGTGCTTGAAGTATGATCAGAGAGGAGACCA CAGGTGTGTCTCC 660  
DB 601 AGGGAGTCTGTGTGTGCTTGAAGTATGATCAGAGAGGAGACCA CAGGTGTGTCTCC 660  
QY 661 CCTTGTGTTAAGCGTGAAGCAGAGAGAGTATGTCACGACATTTCCAAAGGTGGGGG 720  
DB 661 CCTTGTGTTAAGCGTGAAGCAGAGAGAGTATGTCACGACATTTCCAAAGGTGGGGG 720

QY	721	GGTCCGTTGTTCCCGAGATACCTTTTAGTGTAAGGGGCGCTGACATTAAAGTGCAAA	780
Db	721	GGTCCGTTGTTCCCGAGATACCTTTTAGTGTAAGGGGCGCTGACATTAAAGTGCAAA <td>780</td>	780
QY	781	TGAGACCAAGAAAGGAGATGCGCTTCCCAATCTCTCAATCCTTTTAGCGGAGAAAGCT	840
Db	781	TGAGACCAAGAAAGGAGATGCGCTTCCCAATCTCTCAATCCTTTTAGCGGAGAAAGCT	840

QY 841 CAGCTGGATGCCAACAATGTTCCGATGCGTGGGAAGCATGCCGACGTCTCCTCTGCGCTA 900  
 Db 841 CAGCTGGATGCCAACAATGTTCCGATGCGTGGGAAGCATGCCGACGTCTCCTCTGCGCTA 900  
 QY 901 GGGAGCAGCATTTGGGCTTAAAGGACAGGTGGAAAAAATTCCAGCTTTTAACTACATGTTT 960  
 Db 901 GGGAGCAGCATTTGGGCTTAAAGGACAGGTGGAAAAAATTCCAGCTTTTAACTACATGTTT 960

Db	961	TTGTTTAAAGCAATATTTTATTTGGCTACTTATTTGTTAGGACAAAGTGATGGCAT	1020
Qy	1021	TCATATTTATTGTGACCTTTTCAATAAATAGATTTAAAGTAAAAA	1080
Db	1021	TCATATTTATTGTGACCTTTTCAATAAATAGATTTAAAGTAAAAA	1080
Qy	1081	AAAAAGCTGTGGGGGTTACCCGGGGGCCAAAGGGGGGCCCGGGGGGAAATGTGTTTCC	1140
Db	1081	AAAAAGCTGTGGGGGTTACCCGGGGGCCAAAGGGGGGCCCGGGGGGAAATGTGTTTCC	1140
Qy	1141	CGCCCAATTCGCCCCCAATTTTGTGACAAATGAAAGGACACACGGGAATCC	1192
Db	1141	CGCCCAATTCGCCCCCAATTTTGTGACAAATGAAAGGACACACGGGAATCC	1192

RESULT 2	
ABL90057	
ID	ABL90057 standard; cDNA; 1108 BP.
XX	
AC	ABL90057;
XX	
DT	24-MAY-2002 (first entry)

XX Human polynucleotide SEQ ID NO 619.

DE

XX

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; antinflammatory; antitumor; antineoplastic; anticonvulsant; antibacterial; antifungal; antiparasitic; antineurological disease; cancer; immune disorder; cardiovascular disorder; neurological disease; infection; human; secreted protein; gene; ss.

XX

XX  
OS Homo sapiens.  
XX  
FN WC2001,90304-A2.  
XX  
PD 29-NOV-2001.  
XX  
PF 18-MAY-2001; 2001WO-US016450.

XX  
PR 19-MAY-2000; 2000US-020551SP.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Birse CE, Rosen CA;  
XX WPI; 2002-122018/16.  
XX

DR P-FSDB; AB889648.  
XX  
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive,  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders.  
XX  
XX Claim 4; SEQ ID NO 619; 2081BP + Sequence Listing; English.

XX The invention relates to novel genes (AB189449-AB190853) and proteins (AB189040-AB189044) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,

CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patient did not form part of

Sequence 1108 BP; 234 A; 274 C; 324 G; 271 T; 0 U; 5 Other;  
Query Match 89.0%; Score 1061.4; DB 6; Length 1108;  
Best Local Similarity 99.6%; Pred. No. 5.5e-234;  
Matches 1059; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
QY 15 CGCCGGAGTGAAGCGCCGCGAGGAGGAGTGGCTGCGAGTGCAGACAGACATCTACTCAG 74  
DB 1 CGCCGGAGTGAAGCGCCGCGAGGAGGAGTGGCTGCGAGTGCAGACAGACATCTACTCAG 60  
QY 75 CGTGGGTGACCTCTGTGAACATCACTGACTGCAAGCTCCCTCAATTTCTGTGCAAGCC 134  
DB 61 CGTGGGTGACCTCTGTGAACATCACTGACTGCAAGCTCCCTCAATTTCTGTGCAAGCC 120  
QY 135 ATAGAGGAGCCACAGAGGCTGGGAGATGGTGGGATCTTGGCCAAATGGGAAATGTGC 194  
DB 121 ATAGAGGAGCCACAGAGGCTGGGAGATGGTGGGATCTTGGCCAAATGGGAAATGTGC 180  
QY 195 AGAGCAGCAGACCCCGAGTGAAGCACTACAGCCACCAAGAGGTAGCATTTCTCGAC 254  
DB 181 AGAGCAGCAGACCCCGAGTGAAGCACTACAGCCACCAAGAGGTAGCATTTCTCGAC 240  
QY 255 AGAGCTTCTTAAATAGGGGCGCATGTGTCTCCCGAGGGGGTCTTGGCCCGCCAGCAGC 314  
DB 241 AGAGCTTCTTAAATAGGGGCGCATGTGTCTCCCGAGGGGGTCTTGGCCCGCCAGCAGC 300  
QY 315 AGGAGGTGCGAGGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 374  
DB 301 AGGAGGTGCGAGGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 375 TGAACATGGGCTTCCGAGTGGCATCTGCGCAACATGCTGTGAGAGCCGGTACCTCA 434  
DB 361 TGAACATGGGCTTCCGAGTGGCATCTGCGCAACATGCTGTGAGAGCCGGTACCTCA 420  
QY 435 TCCGTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 494  
DB 421 TCCGTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
QY 495 TCTAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 554  
DB 481 TCTAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 555 TTGAGAGGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 614  
DB 541 TTGAGAGGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
QY 615 TGCCTTGAAGGTATGATCAGAGAGGAGCAGCAGGTGTGTGTTCCCTTGTGTTAAG 674  
DB 601 TGCCTTGAAGGTATGATCAGAGAGGAGCAGCAGGTGTGTGTTCCCTTGTGTTAAG 660  
QY 675 GTGAGCAGAGGAGAGCTTATGCTCAGATTTCCAAAGTGTGGGTGCTGCTGCTGCT 734  
DB 661 GTGAGCAGAGGAGAGCTTATGCTCAGATTTCCAAAGTGTGGGTGCTGCTGCTGCT 720  
QY 735 CGAGATACCTTTAGTGTGATGAGGAGCTGATTAAGTGCACAAATCAGAGCAAGAA 794  
DB 721 CGAGATACCTTTAGTGTGATGAGGAGCTGATTAAGTGCACAAATCAGAGCAAGAA 780  
QY 795 CGATGCCCTTCCCAATCTCTCATCTCTTTATGCGAGAAAGATCTAGCTGATGCCAA 854  
DB 781 CGATGCCCTTCCCAATCTCTCATCTCTTTATGCGAGAAAGATCTAGCTGATGCCAA 840  
QY 855 CATGTTCCGATGCTGTGAGAGCATGCCAGCTCTCTGCTGCTGAGAGAGAGACTTG 914  
DB 841 CATGTTCCGATGCTGTGAGAGCATGCCAGCTCTCTGCTGCTGAGAGAGAGACTTG 900  
QY 915 GGCCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 974  
DB 901 GGCCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
QY 975 TATTTTATTTGGCTACTTATTTGTTTGAACAAGTGTGATGCTATTTATTTGTA 1034  
DB 961 TATTTTATTTGGCTACTTATTTGTTTGAACAAGTGTGATGCTATTTATTTGTA 1020

QY 1035 CCTTTCAATAATGATTTAGTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1077  
DB 1021 CCTTTCAATAATGATTTAGTAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1063  
RESULT 3  
AAVS9558  
ID AAVS9558 standard; DNA; 1049 BP.  
XX  
AC AAVS9558;  
XX  
DT 06-JAN-1999 (first entry)  
XX  
DE Human secreted protein gene 48 clone HPCA174.  
XX  
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumor; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX  
OS Homo sapiens.  
XX  
PN MO9839448-A2.  
XX  
PD 11-SEP-1998.  
XX  
PF 06-MAR-1998; 98MO-US004493.  
XX  
XX 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040161P.  
XX 07-MAR-1997; 97US-004162P.  
PR 07-MAR-1997; 97US-0040163P.  
XX 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
XX 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040337P.  
XX 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
XX 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
XX 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
XX 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
XX 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
XX 11-APR-1997; 97US-0043669P.  
PR 11-APR-1997; 97US-0043670P.  
XX 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
XX 11-APR-1997; 97US-0043674P.  
PR 11-APR-1997; 97US-0043675P.  
XX 11-APR-1997; 97US-0043676P.  
PR 11-APR-1997; 97US-0043677P.  
XX 11-APR-1997; 97US-0043678P.  
PR 11-APR-1997; 97US-0043679P.  
XX 11-APR-1997; 97US-0043680P.  
PR 11-APR-1997; 97US-0043681P.  
XX 11-APR-1997; 97US-0043682P.  
PR 11-APR-1997; 97US-0043683P.  
XX 11-APR-1997; 97US-0043684P.  
PR 11-APR-1997; 97US-0043685P.  
XX 11-APR-1997; 97US-0043686P.  
PR 11-APR-1997; 97US-0043687P.  
XX 11-APR-1997; 97US-0043688P.  
PR 11-APR-1997; 97US-0043689P.  
XX 11-APR-1997; 97US-0043690P.  
PR 11-APR-1997; 97US-0043691P.  
XX 11-APR-1997; 97US-0043692P.  
PR 11-APR-1997; 97US-0043693P.





QY 764 CATTAAATGGCAAAATTCAGAGCAGAGAAAGCATGCCCTTCCCAATTCCTCAATCCCT 823  
Db 721 CATTAAATGGCAAAATTCAGAGCAGAGAAAGCATGCCCTTCCCAATTCCTCAATCCCT 780  
QY 824 TTATGGCAGAAAGTCTCAGCTGGATGCCAATGTTCCGATGCCCTGTGGAAGACATGCC 883  
Db 781 TTATGCCAGAAAGTCTCAGCTGGATGCCAATGTTCCGATGCCCTGTGGAAGACATGCC 840  
QY 884 GAGCTCTCTCTGCTTAGGAGAGAGACATTGGCTTAGGGCAGGTGAAAAATTCACA 943  
Db 841 GAGCTCTCTCTGCTTAGGAGAGAGACATTGGCTTAGGGCAGGTGAAAAATTCACA 900  
QY 944 CTTTATAGACCTGTTTTTTTAAATGCTAATTTTATGCTCTTATGTTAGG 1003  
Db 901 CTTTATAGACCTGTTTTTTTAAATGCTAATTTTATGCTCTTATGTTAGG 960  
QY 1004 ACAAGTGTAGTGGCATTTCTATTATTGTCACCTTTTCATAAATAGATTAAAGTAAAA 1063  
Db 961 ACAAGTGTAGTGGCATTTCTATTATTGTCACCTTTTCATAAATAGATTAAAGTAAAA 1020  
QY 1064 AAAAAAAAAAAAAA 1077  
Db 1021 AAAAAAAAAAAAAA 1034

RESULT 4  
ABS73545  
ID ABS73545 standard; cDNA; 1049 BP.  
XX  
AC ABS73545;  
XX  
DT 15-JAN-2003 (first entry)  
XX  
DE Human cDNA #1 for novel secreted protein gene 48.  
XX  
KW Human; ss; gene; secreted protein; autoimmune disease; chemotaxis;  
KW rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;  
KW liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;  
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis; sunburn;  
KW nervous system disorders; Alzheimer's disease; infection;  
KW ocular disorder; corneal infection; wound healing; tissue regeneration;  
KW epithelial cell proliferation; organ transplantation; food additive;  
KW preservative; nutritional.  
XX  
OS Homo sapiens.  
XX  
PN US6420526-B1.  
PD 16-JUL-2002.  
XX  
PF 08-SEP-1998; 98US-00149476.  
XX  
PR 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040161P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040626P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-0043669P.  
PR 11-APR-1997; 97US-0043670P.

PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043674P.  
PR 23-MAY-1997; 97US-0044792P.  
PR 23-MAY-1997; 97US-0047500P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047502P.  
PR 23-MAY-1997; 97US-0047503P.  
PR 23-MAY-1997; 97US-0047581P.  
PR 23-MAY-1997; 97US-0047582P.  
PR 23-MAY-1997; 97US-0047583P.  
PR 23-MAY-1997; 97US-0047584P.  
PR 23-MAY-1997; 97US-0047585P.  
PR 23-MAY-1997; 97US-0047586P.  
PR 23-MAY-1997; 97US-0047587P.  
PR 23-MAY-1997; 97US-0047588P.  
PR 23-MAY-1997; 97US-0047589P.  
PR 23-MAY-1997; 97US-0047590P.  
PR 23-MAY-1997; 97US-0047592P.  
PR 23-MAY-1997; 97US-0047593P.  
PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047616P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 13-JUN-1997; 97US-0048974P.  
PR 08-JUL-1997; 97US-0049610P.  
PR 16-JUL-1997; 97US-0051928P.  
PR 18-JUL-1997; 97US-0052874P.  
PR 22-AUG-1997; 97US-0055724P.  
PR 22-AUG-1997; 97US-0056309P.  
PR 22-AUG-1997; 97US-0056310P.  
PR 22-AUG-1997; 97US-0056332P.  
PR 22-AUG-1997; 97US-0056336P.  
PR 22-AUG-1997; 97US-0056337P.  
PR 22-AUG-1997; 97US-0056342P.  
PR 22-AUG-1997; 97US-0056642P.  
PR 22-AUG-1997; 97US-0056845P.  
PR 22-AUG-1997; 97US-0056846P.  
PR 22-AUG-1997; 97US-0056847P.  
PR 22-AUG-1997; 97US-0056878P.  
PR 22-AUG-1997; 97US-0056879P.  
PR 22-AUG-1997; 97US-0056880P.  
PR 22-AUG-1997; 97US-0056881P.  
PR 22-AUG-1997; 97US-0056882P.  
PR 22-AUG-1997; 97US-0056884P.  
PR 22-AUG-1997; 97US-0056886P.  
PR 22-AUG-1997; 97US-0056887P.  
PR 22-AUG-1997; 97US-0056888P.  
PR 22-AUG-1997; 97US-0056889P.  
PR 22-AUG-1997; 97US-0056892P.  
PR 22-AUG-1997; 97US-0056893P.  
PR 22-AUG-1997; 97US-0056894P.  
PR 22-AUG-1997; 97US-0056903P.  
PR 22-AUG-1997; 97US-0056908P.  
PR 22-AUG-1997; 97US-0056909P.

PR 22-AUG-1997; 97US-0056910P.  
 PR 22-AUG-1997; 97US-0056911P.  
 PR 05-SEP-1997; 97US-0057650P.  
 PR 05-SEP-1997; 97US-0057669P.  
 PR 05-SEP-1997; 97US-0057761P.  
 PR 12-SEP-1997; 97US-0058785P.  
 PR 02-OCT-1997; 97US-0061060P.  
 PR 06-MAR-1998; 98MO-US004493.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC,  
 PI Bednarik DR, Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM,  
 PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA,  
 PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;

XX WPI: 2002-634796/68.  
 DR P-PSDE; AB935227.

XX New isolated human secreted protein for diagnosing, preventing, treating  
 PT or ameliorating medical conditions and used as a food additive or  
 PT preservative.

PS Example 1; SEQ ID NO 58; 129pp; English.

XX The invention relates to an isolated protein that is one of 186 human  
 CC secreted proteins, given in the specification, encoded by one of 309 cDNA  
 CC sequences also given in the specification. The protein is used in a  
 CC pharmaceutical composition used to prevent, treat or ameliorate a medical  
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 CC chickens or sheep. Disorders which are diagnosed or treated include  
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
 CC infections caused by bacteria, viruses and fungi and ocular disorders  
 CC e.g. corneal infection. The polypeptides can also be used to aid wound  
 CC healing and epithelial cell proliferation, to prevent skin aging due to  
 CC sunburn, to maintain organs before transplantation, for supporting cell  
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
 CC polypeptides can also be used as a food additive or preservative to  
 CC increase or decrease storage capabilities, fat content, lipid, protein,  
 CC carbohydrates, vitamins, minerals, cofactors and other nutritional  
 CC components. The present sequence represents a cDNA derived from a gene  
 CC encoding one of the novel human secreted proteins of the invention. Note:  
 CC This sequence did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docid=642052681

SO Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

Query Match 84.3%; Score 1012; DB 6; Length 1049;

Best Local Similarity 99.3%; Pred. No. 1.2e-222; Mismatches 5; Indels 2; Gaps 1;

Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;  
 QY 46 TCGGCGCTGACAGACAGATCTTACTCAGCGTGGGTCACTCTGTGAACATCACTGACTG 105  
 DB 1 TCGGCGCTGACAGACAGATCTTACTCAGCGTGGGTCACTCTGTGAACATCACTGACTG 60  
 QY 106 CAAGCTTCCTCAATTTCTGTGTGAGCCCATCAGGACCCAGCGCTGGAGAGATGGT 165  
 DB 61 CAAGCTTCCTCAATTTCTGTGTGAGCCCATCAGGACCCAGCGCTGGAGAGATGGT 120  
 QY 166 GCGGATCTGGGCGCAATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 225  
 DB 121 GCGGATCTGGGCGCAATGGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 QY 226 CGAGCCACCAAGAGAGATCTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 285  
 DB 181 CGAGCCACCAAGAGAGATCTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 QY 286 CCCAGGGGGGCTCTGGCCCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345

DB 241 CCAGAGGGGGTCTTGGCCCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
 QY 346 CCCCTCAATGACCTCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 405  
 DB 301 CCCCTCAATGACCTCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
 QY 406 CAACCAATGCTGTGAG 465  
 DB 361 CAACCAATGCTGTGAG 420  
 QY 466 TGTTCGTGAG 525  
 DB 421 TGTTCGTGAG 480  
 QY 526 ACCTCTGAG 585  
 DB 481 ACCTCTGAG 540  
 QY 586 AGCAGGAGATTTTGGAG 645  
 DB 541 AGCAGGAGATTTTGGAG 600  
 QY 646 CAGGTGTGTGTTTCCCTTGTGTGTTAAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 703  
 DB 601 CAGGTGTGTGTTTCCCTTGTGTGTTAAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
 QY 704 TTTCCAAAGT 763  
 DB 661 TTTCCAAAGT 720  
 QY 764 CATTAAAGT 823  
 DB 721 CATTAAAGT 780  
 QY 824 TTAATGCCAGAAATCTCAGT 883  
 DB 781 TTAATGCCAGAAATCTCAGT 840  
 QY 884 GACGCTCTCTGT 943  
 DB 841 GACGCTCTCTGT 900  
 QY 944 CTTTGT 1003  
 DB 901 CTTTGT 960  
 QY 1004 ACAAGT 1063  
 DB 961 ACAAGT 1020  
 QY 1064 AAAAAAAAAAAAAA 1077  
 DB 1021 AAAAAAAAAAAAAA 1034

RESULT 5  
 ID ACD82688 standard; cDNA, 1049 BP.  
 XX ACD82688;  
 XX 22-SEP-2003 (first entry)

CDNA sequence #48 containing coding region of a human secreted protein.

Human; secreted protein; hyperproliferative disorder; leukaemia;  
 breast cancer; wound; reproductive disorder; blood-related disorder;  
 haemophilia; thrombocytopenia; immunodeficiency; thymic hypoplasia;  
 Wiskott-Aldrich syndrome; autoimmune disorder; multiple sclerosis;  
 graft-versus-host disease; Hashimoto's thyroiditis; allergy; asthma;  
 viral infection; bacterial infection; fungal infection; AIDS; sepsis;  
 renal disorder; kidney failure; cardiovascular disorder; cytostatic;  
 angina pectoris; cerebral ischaemia; congenital heart defect;

PR	13-JUN-1997;	97US-0049610P.
PR	08-JUL-1997;	97US-0051926P.
PR	16-JUL-1997;	97US-0052874P.
PR	18-AUG-1997;	97US-0055724P.
PR	22-AUG-1997;	97US-0056630P.
PR	22-AUG-1997;	97US-0056631P.
PR	22-AUG-1997;	97US-0056632P.
PR	22-AUG-1997;	97US-0056633P.
PR	22-AUG-1997;	97US-0056637P.
PR	22-AUG-1997;	97US-0056642P.
PR	22-AUG-1997;	97US-0056645P.
PR	22-AUG-1997;	97US-0056649P.
PR	22-AUG-1997;	97US-0056672P.
PR	22-AUG-1997;	97US-0056874P.
PR	22-AUG-1997;	97US-0056875P.
PR	22-AUG-1997;	97US-0056876P.
PR	22-AUG-1997;	97US-0056877P.
PR	22-AUG-1997;	97US-0056879P.
PR	22-AUG-1997;	97US-0056882P.
PR	22-AUG-1997;	97US-0056884P.
PR	22-AUG-1997;	97US-0056886P.
PR	22-AUG-1997;	97US-0056887P.
PR	22-AUG-1997;	97US-0056889P.
PR	22-AUG-1997;	97US-0056892P.
PR	22-AUG-1997;	97US-0056893P.
PR	22-AUG-1997;	97US-0056894P.
PR	22-AUG-1997;	97US-0056903P.
PR	22-AUG-1997;	97US-0056908P.
PR	22-AUG-1997;	97US-0056909P.
PR	22-AUG-1997;	97US-0056910P.
PR	22-AUG-1997;	97US-0056911P.
PR	05-SEP-1997;	97US-0057650P.
PR	05-SEP-1997;	97US-0057659P.
PR	05-SEP-1997;	97US-0057761P.
PR	12-SEP-1997;	97US-0058785P.
PR	09-OCT-1997;	97US-0061660P.
PR	06-MAR-1998;	98MO-US004493.
PR	08-SEP-1998;	98US-00149476.
PR	17-MAR-2000;	2000US-019068P.
XX		
XX	(RUBE/) RUBEN S M.	
PA	(ROSE/) ROSEN C A.	
PA	(SOPP/) SOPPET D R.	
PA	(CART/) CARTER K C.	
PA	(BEDN/) BEDNARIK D P.	
PA	(ENDR/) ENDRESS G A.	
PA	(YUGG/) YU G.	
PA	(NIJ/) NI J.	
PA	(FENG/) FENG P.	
PA	(YOUN/) YOUNG P E.	
PA	(GREE/) GREENE J M.	
PA	(FERR/) FERRIE A W.	
PA	(DUAN/) DUAN D R.	
PA	(HUJ/) HU J.	
PA	(FLOR/) FLORENCE K A.	
PA	(OLSE/) OLSEN H S.	
PA	(FISC/) FISCHER C L.	
PA	(EBEN/) EBNER R.	
PA	(BREM/) BREWER L A.	
PA	(MOOR/) MOORE P A.	
PA	(SHIY/) SHI Y.	
PA	(LAF/) LAFLEUR D W.	
PA	(LIY/) LI Y.	
PA	(ZENG/) ZENG Z.	
PA	(KYAW/) KYAW H.	
XX		
XX	Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednariak DP	

PI Endress GA, Yu G, Ni J, Peng P, Young PE, Greene JM, Ferris AM,  
PI Duan DR, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R,  
PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H,  
XX  
DR WPI: 2003-521800/49.  
DR P-PSDB; ABO34421.  
XX  
PT New genes and its encoded prostate cancer antigen proteins, useful for  
PT preventing, treating, ameliorating or diagnosing e.g. prostate cancers,  
PT thymic hypoplasia, multiple sclerosis, AIDS, angina pectoris or cerebral  
PT ischemia.  
XX  
PS Claim 4; SEQ ID NO 58; 260bp; English.  
XX  
CC The present invention relates to the isolation of novel human secreted  
CC proteins and the polynucleotide sequences encoding them. The invention  
CC also discloses vectors, host cells, antibodies, and recombinant methods  
CC for producing human secreted proteins. The polypeptide and polynucleotide  
CC sequences for the secreted proteins are useful for preventing, treating,  
CC ameliorating or diagnosing medical conditions such as hyperproliferative  
CC disorders (e.g. leukemia or breast cancers), wounds, reproductive  
CC disorders (e.g. blood-related disorders (e.g. hemophilia or thrombocytopenia), immunodeficiencies (e.g. Wiskott-Aldrich syndrome or  
CC thymic hypoplasia), autoimmune disorders (e.g. graft-versus-host disease,  
CC multiple sclerosis or Hashimoto's thyroiditis), allergies (e.g. asthma),  
CC viral or bacterial or fungal infections (e.g. AIDS or sepsis), renal  
CC disorders (e.g. kidney failure), cardiovascular disorders (e.g. angina  
CC pectoris, cerebral ischemia or congenital heart defects), respiratory  
CC disorders, neurological disorders (e.g. Alzheimer's disease or  
CC Parkinson's disease), and inflammations (e.g. Crohn's disease). The  
CC polynucleotide or polypeptide may also be used as vaccine adjuvants.  
CC ACB82441:ACB82950 encode human secreted proteins or their fragments.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site at [seqdata.uspto.gov/psipsidentry.html](http://seqdata.uspto.gov/psipsidentry.html)  
XX  
SQ Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

Query March 84.9%; Score 1012; DB 8; Length 1049;  
Best Local Similarity 99.3%; Pred. No. 1,2e-222;  
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 46 TCGGCGCTGAGACACAGCATCTTACAGCGTGGTCACTTGTGAACATCACTGACTG 105  
DB 1 TCGGCGCTGAGACACAGCATCTTACAGCGTGGTCACTTGTGAACATCACTGACTG 60  
QY 106 CAAGCCCTCCCTCAATTTCTGTGAGAGCCATCAAGGACCCAGCGCTGGAGATGGT 165  
DB 61 CAAGCCCTCCCTCAATTTCTGTGAGAGCCATCAAGGACCCAGCGCTGGAGATGGT 120  
QY 166 GCGGATCTTGGCCCAATGGGAAAATCGTGACAGACGACCCCGAGTAGAGACCACTAC 225  
DB 121 GCGGATCTTGGCCCAATGGGAAAATCGTGACAGACGACCCCGAGTAGAGACCACTAC 180  
QY 226 CCGAGCCCAAGAGGTGATCTTCTGTGACAGAGCTTCTTCAATAGGGGCCATGTGCTCC 285  
DB 181 CCGAGCCCAAGAGGTGATCTTCTGTGACAGAGCTTCTTCAATAGGGGCCATGTGCTCC 240  
QY 286 CCGAGGGGCTCTGAGCCCGCCAGCAGCAGCAGAGGTGCCAGGCTGGTCTGCTAGTC 345  
DB 241 CCGAGGGGCTCTGAGCCCGCCAGCAGCAGCAGAGGTGCCAGGCTGGTCTGCTAGTC 300  
QY 346 CCGCTTCAATGACTCAACCGGACGCTGTGAACATGGGCTTCCGCAAGTGGATCTGG 405  
DB 301 CCGCTTCAATGACTCAACCGGACGCTGTGAACATGGGCTTCCGCAAGTGGATCTGG 360  
QY 406 CAACCATGCTGTGAGCCGGTGAACCTTCATCTGCTCTTCTCTCATGATGCTTGG 465  
DB 361 CAACCATGCTGTGAGCCGGTGAACCTTCATCTGCTCTTCTCTCATGATGCTTGG 420  
QY 466 TGTTCGTGCTCTCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 525

QY 526 ACCCTGAGGGCTGATATGAGGGGTGGCTTTGTTGAGAGGAGCACTTGCTGGGCTTGGTGGAG 585  
DB 481 ACCCTGAGGGCTGATATGAGGGGTGGCTTTGTTGAGAGGAGCACTTGCTGGGCTTGGTGGAG 540  
QY 586 AGCAGGATATTTGAGAGGGATCTGCTGCTGCTTGAAGGTATGATCAGAGGGAGCA 645  
DB 541 AGCAGGATATTTGAGAGGGATCTGCTGCTGCTTGAAGGTATGATCAGAGGGAGCA 600  
QY 646 CAGGTGTGTGTTCCCTTGTGTGTTAAGCTGAGGACAGAGGAGAGCTTGTGCTC--AGCA 703  
DB 601 CAGGTGTGTGTTCCCTTGTGTGTTAAGCTGAGGACAGAGGAGAGCTTGTGCTCAGCAT 660  
QY 704 TTCCCAAGTGTGAGTGGGCTCCGTTGTTCCGAGATCTTTAAGTGTATGAGGCTG 763  
DB 661 TTCCCAAGTGTGAGTGGGCTCCGTTGTTCCGAGATCTTTAAGTGTATGAGGCTG 720  
QY 764 CATTAAGTGACAAATTCAGAGCAGAGAAAGCATGCTTCCCATTTCTCATCTT 823  
DB 721 CATTAAGTGACAAATTCAGAGCAGAGAAAGCATGCTTCCCATTTCTCATCTT 780  
QY 824 TTATGCCGAGAGATCTCAGCTGATGCAACATGTTCCGATGCTGTGAGACATGCC 883  
DB 781 TTATGCCGAGAGATCTCAGCTGATGCAACATGTTCCGATGCTGTGAGACATGCC 840  
QY 884 GAGCTCTCTCTGCTGCTGAGGACAGGACTTGAGGACAGGAGAGAGAGAGAGAGAG 943  
DB 841 GAGCTCTCTCTGCTGCTGAGGACAGGACTTGAGGACAGGAGAGAGAGAGAGAGAG 900  
QY 944 CTTTGTGACCTGTTTGTGTTAAGTATATTTTATGCTATGTTAGT 1003  
DB 901 CTTTGTGACCTGTTTGTGTTAAGTATATTTTATGCTATGTTAGT 960  
QY 1004 ACAAGTGTGATGATCTTATTTATGTAACCTTTCAATTAATATGATTAAGTAA 1063  
DB 961 ACAAGTGTGATGATCTTATTTATGTAACCTTTCAATTAATATGATTAAGTAA 1020  
QY 1064 AAAAAAAAAAAAAA 1077  
DB 1021 AAAAAAAAAAAAAA 1034

RESULT 6  
AB211212  
ID AB211212 standard; cDNA; 562 BP.  
XX  
XX AB211212;  
DT 20-JAN-2003 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 94.  
XX  
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
KW arthritis; cytoskeletal; immunomodulator; neotropic; neuroprotective;  
KW antiparkinsonian; antididiabetic; immunosuppressive; dermatological;  
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
KW antiarthritic; gene; ss.  
OS  
PN Homo sapiens.  
PN MO200270539-A2.  
PD 12-SEP-2002.  
PF 05-MAR-2002; 2002MO-US005095.  
PR 05-MAR-2001; 2001US-0079451.  
XX  
XX (INVT) (INVT) INC

556 TGAGAGGACTTGCTGGGCCTT 577  
|||||  
541 TGAGAGGACTTGCTGGGCCTT 562

Db

RESULT 7  
ID AAS64820 standard; cDNA; 1165 BP.  
XX AC AAS64820;  
XX DT 13-FEB-2002 (first entry)  
DE DNA encoding novel human diagnostic protein #624.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KM food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX OS Homo sapiens.  
XX FN WC200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001MO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HYSE-) HYSEQ INC.  
XX PT Drmacac RT, Liu C, Tang YT;  
XX DR WPI; 2001-639362/73.  
XX DR P-PADB; ABC00633.  
XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

Claim 1; SEQ ID NO 624; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polynucleotide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Seq Sequence 1165 BP; 247 A; 279 C; 342 G; 297 T; 0 U; 0 Other;

Query Match 38.2%; Score 455.4; DB 5; Length 1165;  
Best Local Similarity 79.9%; Pred. No. 1.2e-94;  
Matches 781; Conservative 0; Mismatches 51; Indels 145; Gaps 15

119 ATTTCGTGACCAACCATTGAGGACCCACAGCGCC--TGAGAGATGTCGCCATCTTTG 176

```
Db 308 AATGACAGTGCAGCCATCAGGACCCACAGCGCCCTGGAGAGATGTGCGGATCTTGG 367
Qy 177 CCAATGGGGAATCGTSCAGAGAC-GACGACCCCCGAGTGAAGACCACTAACCCAGCACCA 235
Db 368 CCAATGGGGAATGTGTGACAGACAGACGCCCGAGTGAAGACCACTAACCCAGCACCA 427
Qy 236 AGAGTGAATCTTCGACAGAGC-TTCTTAATAGGAGGCGCATGTGTCTCCCAAGGAGG 294
Db 428 AGAGTGAATCTTCGACAGAGC-TTCTTAATAGGAGGCGCATGTGTCTCCCAAGGAGG 487
Qy 295 TCC-TGGCCCCCGCAGACAGAGS---CAGGTGCAGGCGTGGGTGCTCAG-TCCCCC 349
Db 488 TCCGTGCCCCCGCAGACAGAGS---CAGGTGCAGGCGTGGGTGCTGCTCAGTTCCCC 547
Qy 350 TTCAATGACC-TCAACCGGAGAGTGTG-AAATGAGGCTTTCGCACT--GGCATCTGG 405
Db 548 TTCAATGACCTTAAACCGGAGAGTGTGAACAATGGCTTTCGCACTTGCATCTTGG 607
Qy 406 CAACCATGCTGTGAGCGCGGTGACCTCAATCC-----TGCTCTCTTCTCCGCTCATGATG 460
Db 608 CAACCATGCTGTGAGCGCGGTGAAGATCACTGCTCTCTCTCTCTCTCTCATGATG 667
Qy 461 CTGTGTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 519
Db 668 CTGTGTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 727
Qy 520 GCGGTGACCTCTGAGGCTGTATAGGAGTGGCTTGTGAGAGGACTTGTGCGCTTGG 579
Db 728 GCGGTGACCTCTGAGGCTGTATAGGAGTGGCTTGTGAGAGGACTTGTGCGCTTGG 787
Qy 580 TGTAGAGAGAGGCTATTTGAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 634
Db 788 TGTAGAGAGAGGCTATTTGAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 847
Qy 635 AGAGGGAGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 692
Db 848 AGAGGGAGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 907
Qy 693 TTACTCGAGATTTCCAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 752
Db 908 TTACTCGAGATTTCCAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 967
Qy 753 TATGGGCTGTGATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 812
Db 968 TATGGGCTGTGATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 998
Qy 813 TCTCATTCTTTATGTCCGAGAGATCTGAGCTGATGCAATGTTCGATGCTGTG 872
Db 999 ----- 998
Qy 873 GAGACATGCGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 932
Db 999 -----GAGACATGCGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1028
Qy 933 AAAATTCAGATTTTATGACATGTTTGTATATGATATTTTATATGCTTACTT 992
Db 1029 AAAATTCAGATTTTATGACATGTTTGTATATGATATTTTATATGCTTACTT 1088
Qy 993 TATGTGTTAGAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1052
Db 1089 TATGTGTTAGAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1148
Qy 1053 TTAAGTAAAAA 1069
Db 1149 TTAAGTAAAAA 1165
```

```
RESULT 8
AAF82505
ID AAF82505 standard; cDNA; 550 BP.
XX AAF82505;
```

```
DT 18-JUN-2001 (first entry)
XX
XX Human breast tumour protein cDNA 15364.
DE Human; breast cancer; tumour; cytosolic; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX WO200127276-A2.
XX
XX 19-APR-2001.
XX
XX 11-OCT-2000; 2000WO-US028255.
XX
XX 12-OCT-1999; 99US-00417031.
XX
XX (CORI-) CORIXA CORP.
XX
XX Harlocker SL, Dillon DC, Xu J;
XX
XX WPI, 2001-273773/28.
XX
XX New polypeptides encoded by polynucleotide sequences over-expressed in
XX breast tumor tissue are useful to detect, monitor and treat breast
XX cancer.
XX
XX Claim 13; Page 50; 52pp; English.
XX
XX The present sequence encodes a breast tumour-associated protein. It was
XX shown to have at least two-fold overexpression in breast tumour tissue.
XX The invention relates to an isolated polypeptide comprising at least an
XX immunogenic portion of a breast tumor-specific protein, or its variant
XX that retains the ability to react with antigen-specific antisera. The
XX breast tumour polynucleotides, polypeptides and antibodies are useful for
XX inhibiting development of breast cancer. The polynucleotides may be used
XX to design primers and probes for detecting and monitoring breast cancer
XX
XX Sequence 550 BP; 102 A; 163 C; 164 G; 118 T; 0 U; 3 Other;
XX
XX Query Match 38.0%; Score 452.6; DB 4; Length 550;
XX Best Local Similarity 95.4%; Pred. No. 3.9e-94;
XX Matches 519; Conservative 0; Mismatches 17; Indels 8; Gaps 5;
Qy 47 GCGGCTGACAGACAGCATCTACTGACGCTGTACCTCTGTGAACATCACTGACTGC 106
Db 1 GCGGCTGACAGACAGCATCTACTGACGCTGTACCTCTGTGAACATCACTGACTGC 60
Qy 107 AAGCTCTCTCAATTTCTGTGTGACGCCCATAGAGGACCCAGGCGCTGGAGGATGGTG 166
Db 61 AAGCTCTCTCAATTTCTGTGTGACGCCCATAGAGGACCCAGGCGCTGGAGGATGGTG 120
Qy 167 CGGATCTTGCCCAATGTGGGAATGTGTGAGGACGACGACCCCGAGTGTAGGACCACTACC 226
Db 121 CGGATCTTGCCCAATGTGGGAATGTGTGAGGACGACGACCCCGAGTGTAGGACCACTACC 180
Qy 227 CAGCCACCAAGAGTGTGATCTCTGACAGACTTTCTTCAATAGGGGCTATGTGTCTCC 286
Db 181 CAGCCACCAAGAGTGTGATCTCTGACAGACTTTCTTCAATAGGGGCTATGTGTCTCC 240
Qy 287 CAGGGGGTCTGTGGCCCCCGCAGACAGAGGACAGGTGACAGGTGGGTGCTCACTCC 346
Db 241 CAGGGGGTCTGTGGCCCCCGCAGACAGAGGACAGGTGACAGGTGGGTGCTCACTCC 300
Qy 347 CCTTCAATGACCTCAACCGGAGCTGTGAACATGGGCTTTCCGAGTGTGATCTGGC 406
Db 301 CCTTCAATGACCTCAACCGGAGCTGTGAACATGGGCTTTCCGAGTGTGATCTGGC 360
Qy 407 AACCATGG-TGTGAGGCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 465
Db 361 AACCATGGTGTGAGGCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 419
Qy 466 TGTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 524
```



Tue Aug 24 09:48:33 2004

us-10-001-885-31.rng

Page 12

XX WPI; 2002-471617/50.  
 DR  
 XX  
 PT New ovary specific genes and proteins, useful as a vaccine for treating  
 PT patients with ovarian cancer, or for diagnosing and monitoring the  
 PT presence and metastases of ovarian cancer in a patient.  
 XX  
 XX  
 SS Claim 1; Page 173; 260pp; English.

CC ABBN87799 to ABBN87882 represent human ovary specific nucleic acid (OSNA)  
CC sequences, and ABB79397 to ABB79399 represent ovary specific protein  
CC (OSP) sequences from the present invention. OSNA and OSP sequences have  
CC cytostatic activity, and can be used in vaccine production and gene  
CC therapy. An antibody, that specifically binds to an OSP can be used for  
CC treating a patient with ovarian cancer, particularly for inducing an  
CC immune response against the ovarian cancer cell expressing the OSNA or  
CC OSP. The OSNAs and OSPs can also be used for diagnosing and monitoring  
CC the presence and metastases of ovarian cancer in a patient

XX Sequence 643 BF; 170 A; 139 C; 172 G; 162 T; 0 U; 0 Other;  
XQ

Query Match	36.3%;	Score 432.8;	DB 6;	Length 643;
Best Local Similarity	96.1%;	Pred. No. 1.5e-89;		
Matches 519; Conservative	0;	Mismatches 12;	Indels 9;	Gaps 7;

QY	666	CTTTGCTTAACGGAGGACAGAGGAGAACTCTTGTGTCAG - CATTTCCAAAGTGTGGTG	720
Db	104	CTTTGCTTTAAGCGTGAAGGACAGAGGAGAACTTGTGTCAGAACTTTCCAAAGTGTGGTG	163
QY	721	GGTCCGTTGGTTCCCGAGATACTTTTAAGTGTGTATGSGGCTTCGACATTAAAGTGAC - AAA	779
Db	164	GGTCCCTTGTTCCGAGATACTTTTAAGTGTGTATGSGGCTTCGACATTAAAGTGAC - AAAA	223
QY	780	ATCAGGCAAGAAAGGATGCCCCCTT - CCGAATTCATCAGTCCCTTT - -ATGCCAGAAAG	836
Db	224	ATCAGGCAAGAAAGGATGCCCTTCCCAATTCCTCAATCCTTTTATGGCCGAGAAAG	283
QY	837	ATCTCAGCTGATGCGCAACATGTTCCGATGCTGTGGAAGACATGCCGACGTCTCTGTG	896
Db	284	ATCTCAGCTGATGCGCAACATGTTCCGATGCTGTGGAAGACATGCCGACGTCTCTGTG	343
QY	897	CCGAGGACACAGACTGTGGCTTATGGGACAGGTGGAAGAAAATTCACAGCTTTTATG - CA	954
Db	344	CCGAGGACACAGACTGTGGCTTATGGGACAGGTGGAAGAAAATTCACAGCTTTTATGACAC	403
QY	955	CTGTTTTGTTTAAATGATATAT - TTTAATGGCTACTTATTGTTTAAAGCAAGTGTA	1013
Db	404	TGTAATTTGTTTAAATGATATATCTATATGCTACTTTATGTATAGACAAAGTGTA	463
QY	1014	GTGGCACTCTATTTAT - GTGACCTTTCAATTAATGATTTTAAAGTAAAAA	1072
Db	464	GTGGCACTCTATTTATGTGTGACCTTTTCAATTAATGATTTTAAAGCAAAAAA	523
QY	1073	AAAAAAAAAAAAAGCTGTGGGAGTTACCCGGGAGCAAAAGGAGGAGCCCGGAGGAGATGT	1133
Db	524	AAAAAAAAAAAAAGCTGTGGGAGTTACCCGGGAGCAAAAGGAGGAGCCCGGAGGAGATGT	583
QY	1133	GGTTCCTCCGCCCAATTCGCCCAATTTTGTGCAAAATGAAGAGACACACGAGACCC	1197
Db	584	GGTTCCTCCGCCCAATTCGCCCAATTTTGTGCAAAATGAAGAGACACACGAGACCC	643

RESULT 11  
ABX74673/c  
ID ABX74673 standard; cDNA; 517 BP.

DT 21-MAR-2003 (first entry)  
XX  
DE Human cDNA sequence #128 differentially expressed in CC-RCC types  
XX

KM differential expression profile; aggressive CC-RCC tumour type;  
KM non-aggressive CC-RCC tumour type; clear cell renal carcinoma;  
KM gene expression profiling; tumour tissue; gene; ss.

**Homo sapiens.**

WO200279411-A2.

10-OCT-2002.

29-MAR-2002; 2002WO-US009576.

29-MAR-2001; 2001US-027

(VAND-) VAN ANDEL INST.

Haab B, Rhodes D, Teh BT, Takashi M;

WPI; 2003-040679/03.

New microarray, comprising a matrix of cDNA probe from a set of probes

prognosis of patients with clear cell renal carcinoma.

Claim 1; SEO ID NO 459; 179pp; English.

The present invention relates to microarray comprising a matrix of at least one cDNA probe from a set of probes immobilised to a solid surface in a predetermined order, where a row of pixels corresponds to replicates of one distinct probe from the set. The probes are complementary to nucleic acid sequences that are expressed differentially in aggressive as compared to non-aggressive types of clear cell renal carcinoma (CC-RCC) and that hybridise to the probes under high stringency conditions. The microarray is useful for the prognosis of patients with CC-RCC, wherein aggressive and non-aggressive CC-RCC tumour types are characterised by differential expression profiles of genes that hybridise with one or more probes immobilised on the microarray. The arrays are useful for gene expression profiling of tumour and normal tissues. The present sequence represents a human cDNA sequence differentially expressed in aggressive vs. non-aggressive type CC-RCC phenotypes.

Sequence 517 BP; 166 A; 154 C; 90 G; 107 T; 0 U; 0 Other;

Query Match 31.2%; Score 372.2; DB 7; Length 517;

Best Local Similarity: 80.7%; E-Val: NO: 1.2e-05;  
Matches 514; Conservative 0; Mismatches 3; Indels 120; Gaps 2;

432 CCATCCTGCTCCTCTTCCCTGCTCATGATGCTTGGTGTTCGTGCCCTCCTCCTGTTGCC 491

Db 517 CCATCCTGCTCCTCTTCCCTGCTCATGATGCTTGGTGTTCGTGGCCTCCTCCTGGTGGCC 458

492 TTGCTACCTGGTGTCCCACTGAGTCAGCGGTGACCTCTGAGGGCTGATAGGGTGGT 551

Db 457 TTGCTACCT-GTGTCCCACTGAGTCAGCGGTGACCTCTGAGGGCTGATAGGGGTGGT 399

552 TTGTTGAGAGGACTTGCTGGCCCTTGCTGTGAGAGCAGGCATATTTGAGGGGATCTGG 611

398 TTGTTGAGAGGACTTCCTGGCCCTTGCTGTGAGACGACCATATTTCGAGGGATCTGG 339

612 TGGTGCTTGAAGTATGATCAGAGAGGGACCAAGGTGTGTGTTCCCCCTTGTGTTA 671

338 TCGTCCTTCAAGGTATGATCAGAGAGGGACCAAGAGTGTGTMTCCCTTGTGTTA 279

670 AGGCTGGAGGAGGAGGAGTTCAGCAATTTCCAAAGTGGTGGGTCCGTTGGT 731

378 TCCCTGACCGACGTTAGCAATTTCCAAAGTGAGGTGGGTTCGGT 219

733 TCTCCGCGATTCATTCTTCTCCTGGAGCAACAAATCAGAGCAAGA 791

167

**85]**



Db 166 ----- 167  
 QY 852 CAACATGTCGATGCTGTGAAGACATGCCAGCTCTCTCTGCTAGGAGCAGAC 911  
 Db 166 -----GAGCAGGAC 158  
 QY 912 TTGGGCTTAGGGGAGGTGAAAAAATTCAGACTTTTTCAGCTGTTTGTTCATAG 971  
 Db 157 TTGGGCTTAGGGGAGGTGAAAAAATTCAGACTTTTTCAGCTGTTTGTTCATAG 98  
 QY 972 GTATATTTTATTTGGCTACTTATTTGTTAGCAAGTGTGTAGTGCATTTCATTTATG 1031  
 Db 97 GTATATTTTATTTGGCTACTTATTTGTTAGCAAGTGTGTAGTGCATTTCATTTATG 38  
 QY 1032 TGACCTTTTCAATTAATGATTTAAGTAAAAA 1068  
 Db 37 TGACCTTTTCAATTAATGATTTAAGTAAAAA 1  
 RESULT 12  
 ABX92251  
 ID ABX92251 standard; cDNA, 365 BP.  
 AC ABX92251;  
 XX 08-MAY-2003 (first entry)  
 DT  
 XX Human ovarian specific nucleic acid DEX0310\_65.  
 DS  
 XX Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;  
 KM non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;  
 KM gynaecological.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200292785-A2.  
 PN 21-NOV-2002.  
 PD  
 XX 13-FEB-2002; 2002WO-US022271.  
 PF  
 XX 13-FEB-2001; 2001US-0268290P.  
 PR 15-FEB-2001; 2001US-0268834P.  
 XX  
 PA (DIAD-) DIADEXUS INC.  
 XX  
 PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferey R;  
 PI Sun Y, Liu C;  
 DR P-PSDB; AB061063.  
 XX  
 PT New isolated OSNA nucleic acid and encoded polypeptide, useful for  
 PT identifying, diagnosing, monitoring, staging, imaging and treating  
 PT ovarian cancer and non-cancerous diseases in ovarian tissues.  
 XX  
 PS Claim 1; Page 185-186; 224pp; English.  
 XX  
 XX The invention relates to a new isolated nucleic acid termed ovarian  
 CC specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that  
 CC encodes any of 53 fully defined protein sequences appearing as AB061018-  
 CC AB061070 (termed ovarian specific proteins, OSP); (b) any of 76 fully  
 CC defined nucleotide sequences appearing as ABX92187-ABX92282; or (c) a  
 CC sequence having at least 60% sequence identity to the nucleic acid  
 CC molecule of (a) or (b). Also included are a method for determining the  
 CC presence of an ovary specific nucleic acid (OSNA) in a sample, a vector  
 CC comprising an OSNA, a host cell comprising the vector, an isolated OSP  
 CC polypeptide, an anti-OSP antibody or fragment, a method for determining  
 CC the presence of an ovary specific protein in a sample and a vaccine  
 CC comprising an OSP or OSNA. The methods and compositions of the present  
 CC invention are useful for identifying, diagnosing, monitoring, staging,  
 CC imaging and treating ovarian cancer and non-cancerous disease in ovary  
 CC tissue. The present sequence is an OSNA of the invention

XX SQ Sequence 365 BP; 60 A; 117 C; 108 G; 80 T; 0 U; 0 Other;  
 Query Match 30.5%; Score 363.4; DB 7; Length 365;  
 Best Local Similarity 99.7%; Pred. No. 1,1e-73;  
 Matches 364; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 161 ATGTGCGGATCTTGGCCCAATGAGGAAATGCTGACAGACGACCCCGAGTAGAGAC 220  
 Db 1 ATGTGCGGATCTTGGCCCAATGAGGAAATGCTGACAGACGACCCCGAGTAGAGAC 60  
 QY 221 ACTACCGAGCCAGCAAGAGGTAGCATTCCTGACAGAGCTTCTCAATAGGGGCCATAGT 280  
 Db 61 ACTACCGAGCCAGCAAGAGGTAGCATTCCTGACAGAGCTTCTCAATAGGGGCCATAGT 120  
 QY 281 GCTCCCGCAGGGGGTCTGAGCCGCCGACAGACAGGAGGTGCAGGCTGGTGTCT 340  
 Db 121 GCTCCCGCAGGGGGTCTGAGCCGCCGACAGACAGGAGGTGCAGGCTGGTGTCT 180  
 QY 341 CAGTCCCGCTTCAATGACCTCAACCGGAGCTGTGTGAACATAGGGCTTCCGACATGGAT 400  
 Db 181 CAGTCCCGCTTCAATGACCTCAACCGGAGCTGTGTGAACATAGGGCTTCCGACATGGAT 240  
 QY 401 CTCGGCAACCATGCTGTGAGACCGGTGACCTCCATCCTGCTCTTCTGCTCATGATG 460  
 Db 241 CTCGGCAACCATGCTGTGAGACCGGTGACCTCCATCCTGCTCTTCTGCTCATGATG 300  
 QY 461 CTGGGTGTTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520  
 Db 301 CTGGGTGTTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
 QY 521 CGGTG 525  
 Db 361 CGGTG 365  
 RESULT 13  
 ABX92250  
 ID ABX92250 standard; cDNA, 612 BP.  
 AC ABX92250;  
 XX 08-MAY-2003 (first entry)  
 DT  
 XX Human ovarian specific nucleic acid DEX0310\_64.  
 DS  
 XX Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;  
 KM non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;  
 KM gynaecological.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200292785-A2.  
 PN 21-NOV-2002.  
 PD  
 XX 13-FEB-2002; 2002WO-US022271.  
 PF  
 XX 13-FEB-2001; 2001US-0268290P.  
 PR 15-FEB-2001; 2001US-0268834P.  
 XX  
 PA (DIAD-) DIADEXUS INC.  
 XX  
 PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferey R;  
 PI Sun Y, Liu C;  
 DR P-PSDB; AB061062.  
 XX  
 PT New isolated OSNA nucleic acid and encoded polypeptide, useful for  
 PT identifying, diagnosing, monitoring, staging, imaging and treating  
 PT ovarian cancer and non-cancerous diseases in ovarian tissues.



DT 14-NOV-2002 (first entry)  
XX  
DE Human breast cancer associated coding sequence SEQ ID NO: 38.  
XX  
KM Human; breast specific gene; breast specific protein; breast cancer;  
XX gene therapy; cytostatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200264611-A1.  
XX  
PD 22-AUG-2002.  
XX  
PF 12-FEB-2002; 2002WO-US004197.  
XX  
PR 13-FEB-2001; 2001US-0268292P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferty R;  
PI Sun Y, Liu C;  
XX  
DR WPI; 2002-657582/70.  
XX  
PT New breast specific nucleic acids and proteins, useful for identifying,  
PT diagnosing, monitoring, staging, imaging, and treating breast cancer and  
PT non-cancerous disease states in breast tissue, and in gene therapy.  
XX  
PS Claim 1, Page 179, 367P; English.  
XX  
CC The present invention provides human breast specific coding sequences and  
CC proteins. These can be used in the diagnosis and treatment of breast  
CC cancer and non-cancerous diseases of the breast. The present sequence is  
CC a coding sequence of the invention  
XX  
SQ Sequence 461 BP; 142 A; 84 C; 90 G; 145 T; 0 U; 0 Other;  
  
Query Match 6.3%; Score 75.6; DB 6; Length 461;  
Best Local Similarity 62.3%; Pred. No. 1.7e-07;  
Matches 117; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
  
QY 986 GCTACTTATTGTTAGACAGAGTGTGCACTTCTATTATTGTGACCTTTCAATA 1045  
DB 271 GCTTTTCTCTTTGTGTGCTCAAGGTAGCTGCACTTTTAAATTTGTGAAATATAAAC 330  
  
QY 1046 AATAGATTATAGTAAAAAATAAAAAAAAAAAAAAAAAAGTGTGGGGTTACCGGGG 1105  
DB 331 TTTTGTATTAGAAAAAATAAAAAAAAAAAAAAAAAAGCTTGGGGAAACCGGGG 390  
  
QY 1106 CCAAGGGGGCCCCCGGGGGGAAATGTGTTCTCCGCCCAATTCCCCCAATTGTTG 1165  
DB 391 CCAAAAGCGGTGTCCGGGGGGAATTGTTCTCCGTTCCAAATCCCCCAAAATATCG 450  
  
QY 1166 AAAAA 1171  
DB 451 AGAAGA 456

Search completed: August 22, 2004, 01:19:02  
Job time : 524 secs

Blank

Genoscope Centre National de Séquençage  
BP 191 91006 EVRY cedex - France  
Email: [segrefgenoscope.cns.fr](mailto:segrefgenoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 5629.f For  
more information about this cluster, see  
<http://www.genoscope.cns.fr/>  
<http://cgiblink.cluster.cgiisq@SODFO3CfP12NPiKcluster=5629.f>. Contact  
Feng Liang Email : [fliang@liferesearch.com](mailto:fliang@liferesearch.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

FEATURES  
source  
Paradey Avenue Genoscope sequence ID : CS0DF034CF1NPL.  
Location/Qualifiers  
1. 1059  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DF034Y123"  
/issue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

Query Match  
Best Local Similarity 95.7%; Pred. No. 3,5e-123;  
Matches 987; Conservative 16; Mismatches 24; Indels 4; Gaps 3;

79.6%; Score 949.4; DB 13; Length 1059;

Query 18 CGGAGTGAAGCGCCGAGGCGCCGAGTGCCTGCGACACAGCATCTACTACGGCT 77

Db 1028 CCGGAGTGAAGCGCCGAGGCGCCGAGTGCCTGCGACACAGCATCTACTACGGC- 972

Query 78 GGGTCACTCTGTAACATGACTGACGCAAGCCTCCCTCAATTTCTGGTGACCCCATC 137

Db 971 KGGTCACTCTGTAACATGACTGACGCAAGCCTCCCTCAATTTCTGGTGACCCCATC 912

Query 138 AGGACCCACAGCGCTGGAGAGATGTGGCATTTGGCCAAATGGGAAATGTGCAG 197

Db 911 AGGACCCACAGCGCTGGAGAGATGTGGCATTTGGCCAAATGGGAAATGTGCAG 852

Query 198 ACGACCA-CCCCGAGTGAAGACCACTACCAACCCCAAGAGTGAAGTTCCTGACAG 256

Db 851 ACGACCA-CCCCGAGTGAAGACCACTACCAACCCCAAGAGTGAAGTTCCTGACAG 792

Query 257 AGCTCTTCAATAGGGGCGCATGTGTCTCCCGAGGGGCTCTGGCCCGCGACAGCAG 316

Db 791 AGCTCTTCAATAGGGGCGCATGTGTCTCCCGAGGGGCTCTGGCCCGCGACAGCAG 732

Query 317 GCAAGTCCAGGCTGGTGTCTGCTCAAGTCCCTTCAATGACTCAACCGGACGCTGGTG 376

Db 731 GCAAGTCCAGGCTGGTGTCTGCTCAAGTCCCTTCAATGACTCAACCGGACGCTGGTG 672

Query 377 AACATGGGCTTCCGAGTGGGATCTCGGCAACATGCTGGAGCGGAGACCTTCATC 436

Db 671 AACATGGGCTTCCGAGTGGGATCTCGGCAACATGCTGGAGCGGAGACCTTCATC 612

Query 437 CTGCTCTCTTCTGCTCATGATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 496

Db 611 CTGCTCTCTTCTGCTCATGATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 552

Query 497 TACCTGGTGTCCACCTGAGTCAAGCGGTGACCTCTGAGGGGCTGATAGGGGTTGGT 556

Db 551 TACCTGGTGTCCACCTGAGTCAAGCGGTGACCTCTGAGGGGCTGATAGGGGTTGGT 492

Query 557 GAGAGGACTTGTGTGGCCCTGTGTGTGAAGCAAGGATATTTGAGAGGATCTGTGTGTG 616

Db 491 GAGAGGACTTGTGTGGCCCTGTGTGTGAAGCAAGGATATTTGAGAGGATCTGTGTGTG 432

Query 617 CTTTGAAGGATATGATCAGAGAGGAGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 676

Db 431 CTTTGAAGGATATGATCAGAGAGGAGACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 372

Query 677 GAGGACAGAGGAGACCTTGT 736

Db 371 GAGGACAGAGGAGACCTTGT 312

Query 737 AGATATCTTTAGT 796

Db 311 AGATATCTTTAGT 252

Query 797 ATGCCCTTCCCAATTTCTCTCATGCTTTTATGCCGAGAAATCTCAGCTGATGCCACCA 856

Db 251 ATGCCCTTCCCAATTTCTCTCATGCTTTTATGCCGAGAAATCTCAGCTGATGCCACCA 192

Query 857 TGTTCGATGCTCTGAGAGACATGCGACGCTCTCTCTGCTGAGGAGACGACTTGGG 916

Db 191 TGTTCGATGCTCTGAGAGACATGCGACGCTCTCTCTGCTGAGGAGACGACTTGGG 132

Query 917 CTTAGGCGAGGTGAAAAAATTCAGACTTTTGTAGCACTTTTGTGTGTGTGTGTGTGTGTGTGT 976

Db 131 CTTAGGCGAGGTGAAAAAATTCAGACTTTTGTAGCACTTTTGTGTGTGTGTGTGTGTGTGTGT 72

Query 977 TTTTATGAGTCTTATTTTGT 1036

Db 71 TTTTATGAGTCTTATTTTGT 12

Query 1037 TTTTCAATGAA 1047

Db 11 CTTTCAATGAA 1

RESULT 2  
B1522220/c  
LOCUS  
DEFINITION  
6030811771 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5220570 3',  
mRNA sequence.  
B1522220  
VERSION  
B1522220.1 GI:15347012  
EST.  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
NIH-MGC http://mgi.nci.nih.gov/  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-rcmail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LML at:  
http://image.llnl.gov  
Plate: LML1554 row: 1 column: 19  
High quality sequence start: 31.  
High quality sequence stop: 841.  
Location/Qualifiers  
1. 1027  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5220570"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_120"  
/note="Organ: pooled pancreas and spleen; Vector:  
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
source anonymous pool of spleen and pancreas from 28 yo  
male. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.5 kb, insert size range 1-2.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 025. Note: this is a NIH\_MGC Library."

FEATURES  
source  
1. 1027  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5220570"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_120"  
/note="Organ: pooled pancreas and spleen; Vector:  
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
source anonymous pool of spleen and pancreas from 28 yo  
male. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.5 kb, insert size range 1-2.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 025. Note: this is a NIH\_MGC Library."

Query Match  
Best Local Similarity 60.4%; Score 719.6; DB 12; Length 1027;  
Matches 909; Conservative 0; Mismatches 84; Indels 19; Gaps 12;

Query 3 ACCGCTGAGCCCGCGGAGTGAAGCGCCGAGGCGCCGAGTGTGGCTGTGAGACACA 62



QY 502 GGTGTCCACCTGATGATGAGGCTCTGAGGCTGATAGGGGTGGGTTTGTAGAG 561  
Db 481 GGTGTCCACCTGATGATGAGGCTCTGAGGCTGATAGGGGTGGGTTTGTAGAG 539  
QY 562 GGAATCTGAGGCTGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 621  
Db 540 GGAATCTGAGGCTGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 599  
QY 622 AAGGTATGATGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 681  
Db 600 AAGGTATGATGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 659  
QY 682 AAGGTATGATGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 741  
Db 660 AAGGTATGATGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 716  
QY 742 CTCTTGTGATGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 801  
Db 717 CTCTTGTGATGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 773  
QY 802 CTCTTGTGATGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 861  
Db 774 CTCTTGTGATGATGATGAGGCTGATAGGGGTGGGTTTGTAGAG 833  
QY 862 C 862  
Db 834 C 834

## ORIGIN

Query Match 58.4%; Score 696.2; DB 12; Length 708;  
Best Local Similarity 99.6%; Pred. No. 8.3e-88;  
Matches 698; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Choroid"  
/dev stage="fetal and adult"  
/lab host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone lib="UI-E-EJ0"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site 1: Bcl I; Site 2: Not I;  
UI-E-EJ0 is a subcloned cDNA library constructed  
according to Bonaldo, Lennon, and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT73-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are: fetal eyes,  
AGATTCACGA; lens, CGATTAGCA; eye anterior segment,  
AATCCGACG; optic nerve, CGATTAGCA; retina, CCGG;  
foveal and macular, GTCC; RPE and choroid, ACTTA. This  
library was created for the program, Gene Discovery in the  
Visual System, supported by National Eye Institute (NEI).  
TAG\_Lib=UI-E-EJ0  
TAG\_SEQ=GTCC"

RESULT 4  
LOCUS BM675290/c 708 bp mRNA linear EST 27-FEB-2002  
DEFINITION UI-E-EJ0-ahr-n-23-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone  
BM675290  
VERSION BM675290.1 GI:18985188  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
EXTRACT Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 708)  
AUTHORS Bonaldo, M.F., Lennon, G., and Soares, M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: Bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 1-42, >AT-richlow\_complexity (matched complement)  
Seq primer: M13 Forward  
POLYA=yes.

## FEATURES

Location/Qualifiers  
1..708  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ0-ahr-n-23-0-UI"  
/tissue type="fetal eyes, lens, eye anterior segment,  
optic nerve, retina, fovea, and macular RPE and



```

source
1..937
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5534295"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_1fb="NTH MGC 71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb. "
```







QY	190	CGTGAAGAGCAACGACCCCGAGTGAAGACCACTAACCAAGCAACAAAGATGACATTC	249
Db	182	CGTGAAGAGTACGACCCCGAGTGAAGACCACTAACCAAGCAACAAAGATGACATTC	241
QY	250	TGCACAGACTTCTTCAATAGAGGACCATAGTGTCTCCCAAGAGGATCTTGGCCCGCA	309
Db	242	TGCACAGACTTCTTCAATAGAGGACCATAGTGTCTCCCAAGAGGATCTTGGCCCGCA	301
QY	310	GCAGCAGGAGGAGGACAGGCTGGGCTCATGTCCCGTTCAATGACTCAACCGCA	369
Db	302	GCAGCAGGAGGAGTCCAGGCTGGGCTCATGTCCCGTTCAATGACTCAACCGCA	361
QY	370	GCTGTGAACATAGGCTTTCCGACATGTGCATCTCGACACACATGCTGTGAGCCGATAC	429
Db	362	GCTGTGAACATAGGCTTTCCGACATGTGCATCTCGACACACATGCTGTGAGCCGATAC	421
QY	430	CTCATCTCGCTCTCTCTCTCTCATATGCTTGTTGCTGAGCCCTCCTCGTGTGG	489
Db	422	CTCATCTCGCTCTCTCTCTCTCATATGCTTGTTGCTGAGCCCTCCTCGTGTGG	481
QY	490	CCTTGTCTACCTGTGTGCCACTGATGACGGGTGAACCTTGAAGGCTGATAGGGGTGG	549
Db	482	CCTTGTCTACCTGTGTGCCACTGATGACGGGTGAACCTTGAAGGCTGATAGGGGTGG	541
QY	550	GTTTGTGAAGAGGACCTTGCTGGGCTTGTTGTTGATGAGACAGGCATA-TTTGAGAGGGATC	608
Db	542	GTTTGTGAAGAGGACTTGTCTGGGCTTGTTGTTGATGAGACAGGCATA-TTTGAGAGGGATC	600
QY	609	TGTTGTGTCCTTGAAGATGATCAGAGAGGAGACCAAGGTGTGT-TTCCCTTTGT	667
Db	601	TGTTGTGTCCTTGAAGATGATCAGAGAGGAGACCAAGGAGTGTGTTTCCCTTTGT	660
QY	668	GTTAACTGAGAGGACAGGAGACGTTATGTCACAGATTTCCAAAGTGTGGTG 720	
Db	661	GTTAACTGAGAGCA-AGGAGACCTTATGTCACAGATTTCAAGTGTGGTG 712	

```

RESULT 11
BI193763
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BI193763 713 bp mRNA linear EST 10-JUL-2001
602948418r1 NIH_MGC_42 Homo sapiens CDNA clone IMAGE:5091412 5',
mRNA sequence.
BI193763
BI193763.1 GI:14648783
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 713)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: L1CM1855 row: d column: 05
High quality sequence stop: 709.
Location/Qualifiers
1. 713
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5091412"
/tissue_type="epithelial carcinoma cell line"
/lab_host="DH10B (phage-resistant)"

```

```

/clone_lib="NIH_MGC_42"
/notes="Organ: pancreas; Vector: pDRB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCGACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

```

Query Match	54.5%;	Score 649.2;	DB 12,	Length 713;
Best Local Similarity	98.2%;	Pred. No. 2.9e-81;		
Matches 699;	Conservative	0;	Mismatches 8;	Indels 5;
				Gaps 4;

[illegible]

SOURCE ORGANISM	REFERENCE
Homo sapiens (human)	1 (bases 1 to 1104)
Homo sapiens	1 (bases 1 to 1104)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1 (bases 1 to 1104)
NIH-MGC	NIH-MGC <a href="http://mgi.nci.nih.gov/">http://mgi.nci.nih.gov/</a>
National Institutes of Health	National Institutes of Health
Unpublished (1999)	Unpublished (1999)
Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.

FEATURES

SOURCE

ORIGIN

	Query Match	Similarity	52.8%	Score	629.8	DB	12	Length	1104
	Best Local	Similarity	90.8%	Pred.	1.1e-78				
	Matches	715	Conservative	0	Mismatches	67	Indels	5	Gaps
QY	10	GGAGCCCGCCGGAGGTGACAGCCCGCCGAGGCCCGGAGTGTGGCGCTTGACAGACACAGCATCTTA	69						
Db	1	GGAGCCCGCCGGAGGTGAGAGCGCCCGCCGAGGCCCGGAGTGTGGCGCTTGACAGACAGCATCTTA	60						
QY	70	CTCAGCGTGGGTGACCTCTGTGAACATCACTGACCTGCAAGCTCCCTCAATTTCTGTGTC	129						
Db	61	CTCAGCGTGGGTGACCTCTGTGAACATCACTGACCTGCAAGCTCCCTCAATTTCTGTGTC	120						
QY	130	AGCCCATCAAGGAGCCCAAGCGCTGTGGAGAGTGTGCGGATCTTGTGCGAATGGGGAAT	189						
Db	121	AGCCCATCAAGGAGCCCAAGCGCTGTGGAGAGTGTGCGGATCTTGTGCGAATGGGGAAT	180						
QY	190	CGTGCAGAGACACAGACCCCGAGGTGAGGAGCACTAACCCAGCTCACCAAGAGGTAGCATTTCC	249						
Db	181	CGTGCAGAGTACAGACCCCGAGGTGAGGAGCACTAACCCAGCTCACCAAGAGGTAGCATTTCC	240						
QY	250	TCGACAGAGCTTCTTCATAGAGGAGCAATGTGTCTCCCGCCAGGGAGTCTTGCCCGCCCA	309						
Db	241	TCGACAGAGCTTCTTCATAGAGGAGCAATGTGTCTCCCGCCAGGGAGTCTTGCCCGCCCA	300						
QY	310	GGAGCAGAGAGGTGCAAGCGTGGGTGCTGCTCACTCCCGCTTAAGACTCAACCGGCA	369						
Db	301	GGAGCAGAGAGGTGCAAGCGTGGGTGCTGCTCACTCCCGCTTAAGACTCAACCGGCA	360						
QY	370	GCTGTGAAACATGGGCTTTCGAGAGTGACATCTGCGAACATGCTGTGGAGCCGATGAC	429						
Db	361	GCTGTGAAACATGGGCTTTCGAGAGTGACATCTGCGAACATGCTGTGGAGCCGATGAC	420						
QY	430	CTCCATCTCTGCTCTCTTCTCTGCTGCTGATGATGCTTGTGGGCTCTCTCTGTGTGG	489						
Db	421	CTCCATCTCTGCTCTCTTCTCTGCTGCTGATGATGCTTGTGGGCTCTCTCTGTGTGG	480						

QY 490 CTTTGTACTACCTGGTGTCCCACTCGATGACCGGTGACTCTAGAGGCTGATATGGGGTGG 549

Db 481 CTTGTGTACTACCTGGTGTCCCACTGAGTCACGCGGTACTCTGAGGGCTGATATGGGGTGG 540

QY 550 GTTTGTAGAGGGACTTGTCTGGCCCTTGTGTGAGAGCAGAGATATTTTGGAGGGATTT 609

Db 541 G-TAGTGAACGAGGACTTGTCTGGGGCCCTGCTGTGAGAGCAGGATATTTGAGGGGATTT 599

QY 610 GGTGTG-TGCTTGAAGGTATGATCA-GAGAGGGGACCAAGGTGTGTGTTTCCCTTTGT 667

Db 600 GGTGTGGCTTTGAAGGTATATCAGAGAGGGGACCAAGGTTGTGTATCCCTTTGT 659

QY 668 GTTTAGCGTGAAGGCGCAGAGGGAGACGTTAGTCCAGCATTTCCAAAGTGTGGTGGGCGGT 727

Db 660 GTTTAGCGTTTAGGCAAGGSAAGGTAAGTCCAGGATTTCAAAGTGTGTGTAGGAAACGT 719

QY 728 TGGTTCGCCAGATACTTT-TAGGTGTATGTGGGCTTGCACTTAASTGGCACAAAATCAGA 785

Db 720 GGGTCCACAGAGACATTTCAACAGCGGACACGGGGCGAGGACTAAGTGGGCCCAAGACTCAA 779

QY 786 GCAAGAA 792

Db 780 CCCGAA 786

[illegible]

## ORIGIN

Query Match 45.4%; Score 541.6; DB 14; Length 566;

Best Local Similarity 98.0%; Pred. No. 3.3e-66; Matches 547; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AAACCGCTGAGACCGCCGGAGTGAACGCGCCGAGGAGGAGTGGCCCTGACAGCA 60  
Db 9 AAACCGCTGAGACCGCCGGAGTGAACGCGCCGAGGAGGAGTGGCCCTGACAGCA 68  
QY 61 CAGCATCTACTCAGCGTGGTCACTCTGTGAACATCAGTCTGCAAGCTTCCCTCAAT 120  
Db 69 CAGCATCTACTCAGCGTGGTCACTCTGTGAACATCAGTCTGCAAGCTTCCCTCAAT 128  
QY 121 TTCTGGTGAAGCCCATCAAGGAGCCCAAGCGCTGGGAGGAGTGTGGATCTTGGACAA 180  
Db 129 TTCTGGTGAAGCCCATCAAGGAGCCCAAGCGCTGGGAGGAGTGTGGATCTTGGACAA 188  
QY 181 TGGGGAATGTGACAGACGACGACCCCGAGTGAAGACCACTACCAAGCCAAAGAG 240  
Db 189 TGGGGAATGTGACAGACGACGACCCCGAGTGAAGACCACTACCAAGCCAAAGAG 248  
QY 241 TAGCATTTCTCCGACAGAGCTTTTCAATAGGGGCAATGGTCTCCCGAGGGGTCTTG 300  
Db 249 TAGCATTTCTCCGACAGAGCTTTTCAATAGGGGCAATGGTCTCCCGAGGGGTCTTG 308  
QY 301 CCCCCGACAGACAGGACAGGCTGCGTGGTGTGCTCACTCCCTTCATGACCT 360  
Db 309 CCCCCGACAGACAGGACAGGCTGCGTGGTGTGCTCACTCCCTTCATGACCT 368  
QY 361 CAACCGGACAGTGTGAACATGGGCTTTCCGACGTGGCATCTGGACCACTGCTGTGA 420  
Db 369 CAACCGGACAGTGTGAACATGGGCTTTCCGACGTGGCATCTGGACCACTGCTGTGA 428  
QY 421 GCGGGTACCTCCATCCCTGCTCTCTGCTCATGATGCTGTGTTGTTGAGCTCCCT 480  
Db 429 GCGGGTACCTCCATCCCTGCTCTCTGCTCATGATGCTGTGTTGTTGAGCTCCCT 488  
QY 481 CTTGGTGGCTTGTCTAAGTGTGTCACCTGAGTCAAGGCTGACCTTGAAGGCTGA 540  
Db 489 CTTGGTGGCTTGTCTAAGTGTGTCACCTGAGTCAAGGCTGACCTTGAAGGCTGA 548  
QY 541 TAGGGGTGGTGTGTGA 558  
Db 549 TAGGGGTGGTGTGTGA 566

RESULT 14  
CA942743/c 541 bp mRNA linear EST 30-DEC-2002  
LOCUS 1166105.X1 HR85 islet Homo sapiens cDNA clone IMAGE:6607449 3',  
DEFINITION mRNA sequence.

ACCESSION CA942743  
VERSION CA942743.1 GI:27431223  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Butleria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 541)  
Melton, D., Brown, J., Kenty, G., Pernut, A., Lee, C., Kaestner, K.,  
Lemikha, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blaisstein, A.,  
Schmitt, A., Theising, B., Rittler, E., Ronko, I., Bennett, J.,  
Cardenas, M., Gibbons, M., McCann, R., Cole, K., Tsagarelis, R.,  
Williams, T., Jackson, Y., and Bowers, Y.

TITLE  
JOURNAL  
COMMENT Endocrine Pancreas Consortium  
Unpublished (2000)  
Other ESTs: 1166105.Y1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu  
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@im.wustl.edu)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 455.  
Location/Qualifiers

#### FEATURES

source  
1..541  
/organism="Homo sapiens"  
/mol\_type="rRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6607449"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site: 1:  
Notif. Site 2: XhoI; cDNA made by oligo-dT priming.  
Size-selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning. MD,  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Pernut Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

#### ORIGIN

Query Match 45.0%; Score 536.2; DB 14; Length 541;  
Best Local Similarity 99.4%; Pred. No. 1.9e-65;  
Matches 538; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 531 TGAGGGCTGATAGGGGAGGTTGTTGAGAGGAGCTTGTGGGCTGTGTGAGAGAG 590  
Db 541 TGAGGGCTGATAGGGGAGGTTGTTGAGAGGAGCTTGTGGGCTGTGTGAGAGAG 482  
QY 591 GCATATTGAGAGGAGATCTGTGTGCTTGAAGATGATCAGAGAGGAGACAGAGT 650  
Db 481 GCATATTGAGAGGAGATCTGTGTGCTTGAAGATGATCAGAGAGGAGACAGAGT 422  
QY 651 GTGTGTTTCCCTTGTGTTAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 710  
Db 421 GTGTGTTTCCCTTGTGTTAAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362  
QY 711 AGTGTGGTGGGTCGTTGTTCCGAGATCTTTAGTGGTATGGGGCTGTGATTAAG 770  
Db 361 AGTGTGGTGGGTCGTTGTTCCGAGATCTTTAGTGGTATGGGGCTGTGATTAAG 302  
QY 771 TGGCAAAATCAGAGCAAGAAAGGAGTGCCTTCCCAATTCCTCAATCCTTTATGCC 830  
Db 301 TGGCAAAATCAGAGCAAGAAAGGAGTGCCTTCCCAATTCCTCAATCCTTTATGCC 242  
QY 831 GAGAAGATCTCAGCTGAGTGCACATGTCGATGCTGTGAAAGACATGCCAGCTCT 890  
Db 241 GAGAAGATCTCAGCTGAGTGCACATGTCGATGCTGTGAAAGACATGCCAGCTCT 182  
QY 891 CCTGTGCTTGAAGGAGACGACTTGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 950  
Db 181 CCTGTGCTTGAAGGAGACGACTTGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAG 122  
QY 951 AGCATGTTTGTGTTTAAAGTATTTTATTTAGGCTTATTTGTTTGGCAAGTGG 1010  
Db 121 AGCATGTTTGTGTTTAAAGTATTTTATTTAGGCTTATTTGTTTGGCAAGTGG 62  
QY 1011 GTAGTGATTTATTTATTTATTTGACCTTTTCAATTAATAGATTTAATTAATTAAT 1070  
Db 61 GTAGTGATTTATTTATTTATTTGACCTTTTCAATTAATAGATTTAATTAATTAAT 2  
QY 1071 A 1071  
Db 1 A 1

RESULT 15

Tue Aug 24 09:48:35 2004

us-10-001-885-31.rst

Page 12

BU733254/c  
LOCUS BU733254 533 bp mRNA linear EST 09-OCT-2002  
DEFINITION UI-E-CII-afs-m-18-0-UI-s1 UI-E-CII Homo sapiens cDNA clone  
UI-E-CII-afs-m-18-0-UI 3', mRNA sequence.  
ACCESSION BU733254  
VERSION BU733254.1 GI:23659969  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 533)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 1-35, >AT-richlow\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLYA=yes.

FEATURES  
source

Location/Qualifiers  
1..533  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-CII-afs-m-18-0-UI"  
/tissue\_type="RPE and Choroid"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-CII"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
UI-E-CII is a normalized cDNA library containing the  
following tissue(s): RPE and Choroid. The library was  
constructed according to Bonaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a  
Not I site. Double stranded cDNA was ligated to an EcoR I  
adaptor, digested with Not I, and cloned directionally  
into pT73-Pac vector. The oligonucleotide used to prime  
the synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is ACCTA.  
This library was created for the program, Gene Discovery  
in the Visual System, supported by National Eye Institute  
(NEI).  
TAG\_TISSUE=RPE and Choroid  
TAG\_LIB=UI-E-CII  
TAG\_SEQ=ACCTA"

Query Match 44.5%; Score 530.4; DB 13; Length 533;  
Best Local Similarity 99.6%; Pred. No. 1.2e-64;  
Matches 531; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 604 GGATCTGGTGGCTGCTTGAAGATGATCAGAGAGGGGACCAAGGTGTGTTCCTT 663  
Db 473 GGATCTGGTGGCTGCTTGAAGATGATCAGAGAGGGGACCAAGGTGTGTTCCTT 414  
QY 664 TTGTCTTAAGGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 723  
Db 413 TTGTCTTAAGGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354  
QY 724 CCGTTGTTCCCGAGATACCTTTAGTGTATGGGGCTGCTATTAAGTGACAAAATCA 783  
Db 353 CCGTTGTTCCCGAGATACCTTTAGTGTATGGGGCTGCTATTAAGTGACAAAATCA 294  
QY 784 GAGCAGAGAAAGAGATGCCCTTCCCAATTTCTCATCTTTTATGCCGAGAAATCTCAG 843  
Db 233 GAGCAGAGAAAGAGATGCCCTTCCCAATTTCTCATCTTTTATGCCGAGAAATCTCAG 234  
QY 844 CTGATGCCCAACATGTTCCGATGCTGTGGAAGACATGCCAGCTCTCTGCTAGGG 903  
Db 233 CTGATGCCCAACATGTTCCGATGCTGTGGAAGACATGCCAGCTCTCTGCTAGGG 174  
QY 904 AGCAGACTTGGGCTTAGGSCAGGTGAGAAAAAATTCAGACTTTTATGACATGTTTG 963  
Db 173 AGCAGACTTGGGCTTAGGSCAGGTGAGAAAAAATTCAGACTTTTATGACATGTTTG 114  
QY 964 TTTTAAATGATATTTTATTTATGCTACTTATTTAGGCAAGTGTGTGATCTT 1023  
Db 113 TTTTAAATGATATTTTATTTATGCTACTTATTTAGGCAAGTGTGTGATCTT 54  
QY 1024 ATTATTTGACCTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1076  
Db 53 ATTATTTGACCTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1

Search completed: August 22, 2004, 03:29:18  
Job time : 3005 secs

QY 544 GGGTGGCTTTGTTGAGAGGAGCTTGCTGGCTTGTGTGTGAGAGAGCATATTTGAGG 603



OM nucleic - nucleic search, using sw model

Run on: August 22, 2004, 00:20:26 ; Search time 117 Seconds

(without alignments)  
5653.859 Million cell updates/sec

Title: US-10-001-885-31

Sequence: 1 aaaccgcctgcagccgcg.....gaagagacacacggaacc 1192

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*
- 2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*
- 3: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*
- 4: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*
- 5: /cgn2\_6/prodata/2/ina/PCTUS COMB.seq.\*
- 6: /cgn2\_6/prodata/2/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1012	84.9	1049	4	US-09-149-476-58
2	473	39.7	483	4	US-09-621-976-864
3	391.8	32.9	519	4	US-09-621-976-3064
4	232.8	19.5	364	4	US-09-621-976-17202
5	47.2	4.0	1441	3	US-08-821-894-63
6	45.8	3.8	1971	2	US-08-882-770-4
7	45.6	3.8	664	4	US-09-504-615-66
8	45.4	3.8	2409	3	US-09-293-322C-8
9	45.4	3.8	2409	4	US-09-839-497A-8
10	45.2	3.8	2485	4	US-09-889-463A-9
11	45.2	3.8	3138	1	US-07-867-106-4
12	45.2	3.8	640681	4	US-09-790-988-1
13	45	3.8	3994	4	US-09-738-946-7
14	44.8	3.8	2381	1	US-08-021-608D-9
15	44.8	3.8	2381	1	US-08-726-160-9
16	44.8	3.8	2381	5	PCT-US94-01782-9
17	44.8	3.8	2384	1	US-08-021-608D-1
18	44.8	3.8	2384	1	US-08-726-160-1
19	44.8	3.8	2384	5	PCT-US94-01782-1
20	44.6	3.7	1466	4	US-08-984-819A-10
21	44.6	3.7	1466	4	US-08-984-919A-12
22	44.6	3.7	1472	3	US-08-781-420-10
23	44.6	3.7	1472	3	US-08-781-420-12
24	44.6	3.7	1472	4	US-08-874-102-10
25	44.6	3.7	1472	4	US-08-874-102-12
26	44.6	3.7	1472	4	US-09-006-595A-10
27	44.6	3.7	1472	4	US-09-006-595A-12

28	44.6	3.7	1875	4	US-08-984-919A-46	Sequence 46, Appl
29	44.6	3.7	1875	4	US-08-984-919A-48	Sequence 48, Appl
30	44.6	3.7	1881	4	US-08-874-102-46	Sequence 46, Appl
31	44.6	3.7	1881	4	US-08-874-102-48	Sequence 48, Appl
32	44.6	3.7	6638	2	US-08-070-301-2	Sequence 2, Appl
33	44.4	3.7	819	4	US-09-328-475C-236	Sequence 236, Appl
34	44.4	3.7	1049	4	US-09-800-729-67	Sequence 67, Appl
35	44.4	3.7	1313	4	US-09-149-476-112	Sequence 112, Appl
36	44.2	3.7	396	4	US-09-640-173-33	Sequence 33, Appl
37	44.2	3.7	396	4	US-09-713-550-33	Sequence 33, Appl
38	44.2	3.7	923	3	US-09-004-731-66	Sequence 66, Appl
39	44.2	3.7	923	3	US-08-749-699-66	Sequence 66, Appl
40	44.2	3.7	923	4	US-09-004-729-66	Sequence 66, Appl
41	44	3.7	154	4	US-09-621-976-7748	Sequence 7748, Ap
42	44	3.7	154	4	US-09-621-976-7782	Sequence 7782, Ap
43	44	3.7	593	4	US-09-904-615-59	Sequence 59, Appl
44	44	3.7	593	4	US-09-904-615-59	Sequence 59, Appl
45	43.8	3.7	923	3	US-08-906-769-144	Sequence 144, App

#### ALIGNMENTS

RESULT 1  
US-09-149-476-58  
Sequence 58, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
FILE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002p1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-05-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23

**Tue Aug 24 09:48:34 2004**

us-10-001-885-31.rni

Page 2

1	EARLIER	APPLICATION NUMBER:	60/047, 500
2	EARLIER	FILING DATE:	1997-05-23
3	EARLIER	APPLICATION NUMBER:	60/047, 587
4	EARLIER	FILING DATE:	1997-05-23
5	EARLIER	APPLICATION NUMBER:	60/047, 492
6	EARLIER	FILING DATE:	1997-05-23
7	EARLIER	APPLICATION NUMBER:	60/047, 598
8	EARLIER	FILING DATE:	1997-05-23
9	EARLIER	APPLICATION NUMBER:	60/047, 613
10	EARLIER	FILING DATE:	1997-05-23
11	EARLIER	APPLICATION NUMBER:	60/047, 582
12	EARLIER	FILING DATE:	1997-05-23
13	EARLIER	APPLICATION NUMBER:	60/047, 596
14	EARLIER	FILING DATE:	1997-05-23
15	EARLIER	APPLICATION NUMBER:	60/047, 612
16	EARLIER	FILING DATE:	1997-05-23
17	EARLIER	APPLICATION NUMBER:	60/047, 632
18	EARLIER	FILING DATE:	1997-05-23
19	EARLIER	APPLICATION NUMBER:	60/047, 601
20	EARLIER	FILING DATE:	1997-05-23
21	EARLIER	APPLICATION NUMBER:	60/043, 580
22	EARLIER	FILING DATE:	1997-04-11
23	EARLIER	APPLICATION NUMBER:	60/043, 568
24	EARLIER	FILING DATE:	1997-04-11
25	EARLIER	APPLICATION NUMBER:	60/043, 314
26	EARLIER	FILING DATE:	1997-04-11
27	EARLIER	APPLICATION NUMBER:	60/043, 569
28	EARLIER	FILING DATE:	1997-04-11
29	EARLIER	APPLICATION NUMBER:	60/043, 311
30	EARLIER	FILING DATE:	1997-04-11
31	EARLIER	APPLICATION NUMBER:	60/043, 671
32	EARLIER	FILING DATE:	1997-04-11
33	EARLIER	APPLICATION NUMBER:	60/043, 674
34	EARLIER	FILING DATE:	1997-04-11
35	EARLIER	APPLICATION NUMBER:	60/043, 669
36	EARLIER	FILING DATE:	1997-04-11
37	EARLIER	APPLICATION NUMBER:	60/043, 312
38	EARLIER	FILING DATE:	1997-04-11
39	EARLIER	APPLICATION NUMBER:	60/043, 313
40	EARLIER	FILING DATE:	1997-04-11
41	EARLIER	APPLICATION NUMBER:	60/043, 672
42	EARLIER	FILING DATE:	1997-04-11
43	EARLIER	APPLICATION NUMBER:	60/043, 315
44	EARLIER	FILING DATE:	1997-04-11
45	EARLIER	APPLICATION NUMBER:	60/048, 974
46	EARLIER	FILING DATE:	1997-06-06
47	EARLIER	APPLICATION NUMBER:	60/056, 886
48	EARLIER	FILING DATE:	1997-08-22
49	EARLIER	APPLICATION NUMBER:	60/056, 877
50	EARLIER	FILING DATE:	1997-08-22
51	EARLIER	APPLICATION NUMBER:	60/056, 889
52	EARLIER	FILING DATE:	1997-08-22
53	EARLIER	APPLICATION NUMBER:	60/056, 893
54	EARLIER	FILING DATE:	1997-08-22
55	EARLIER	APPLICATION NUMBER:	60/056, 630
56	EARLIER	FILING DATE:	1997-08-22
57	EARLIER	APPLICATION NUMBER:	60/056, 878
58	EARLIER	FILING DATE:	1997-08-22
59	EARLIER	APPLICATION NUMBER:	60/056, 662
60	EARLIER	FILING DATE:	1997-08-22
61	EARLIER	APPLICATION NUMBER:	60/056, 872
62	EARLIER	FILING DATE:	1997-08-22
63	EARLIER	APPLICATION NUMBER:	60/056, 882
64	EARLIER	FILING DATE:	1997-08-22
65	EARLIER	APPLICATION NUMBER:	60/056, 637
66	EARLIER	FILING DATE:	1997-08-22
67	EARLIER	APPLICATION NUMBER:	60/056, 903
68	EARLIER	FILING DATE:	1997-08-22
69	EARLIER	APPLICATION NUMBER:	60/056, 888
70	EARLIER	FILING DATE:	1997-08-22
71	EARLIER	APPLICATION NUMBER:	60/056, 879

EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,894
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,911
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,936
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,874
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,910
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,864
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,631
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,845
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,892
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/047,588
EARNER FILING DATE:	1997-05-23
EARNER APPLICATION NUMBER:	60/047,585
EARNER FILING DATE:	1997-05-23
EARNER APPLICATION NUMBER:	60/047,586
EARNER FILING DATE:	1997-05-23
EARNER APPLICATION NUMBER:	60/047,590
EARNER FILING DATE:	1997-05-23
EARNER APPLICATION NUMBER:	60/047,594
EARNER FILING DATE:	1997-05-23
EARNER APPLICATION NUMBER:	60/047,589
EARNER FILING DATE:	1997-05-23
EARNER APPLICATION NUMBER:	60/047,593
EARNER FILING DATE:	1997-05-23
EARNER APPLICATION NUMBER:	60/043,670
EARNER FILING DATE:	1997-04-11
EARNER APPLICATION NUMBER:	60/043,578
EARNER FILING DATE:	1997-04-11
EARNER APPLICATION NUMBER:	60/043,576
EARNER FILING DATE:	1997-04-11
EARNER APPLICATION NUMBER:	60/047,501
EARNER FILING DATE:	1997-05-23
EARNER APPLICATION NUMBER:	60/043,614
EARNER FILING DATE:	1997-05-23
EARNER APPLICATION NUMBER:	60/043,578
EARNER FILING DATE:	1997-04-11
EARNER APPLICATION NUMBER:	60/043,576
EARNER FILING DATE:	1997-04-11
EARNER APPLICATION NUMBER:	60/047,501
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,664
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,876
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,881
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,909
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,875
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,862
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,887
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/056,908
EARNER FILING DATE:	1997-08-22
EARNER APPLICATION NUMBER:	60/048,964
EARNER FILING DATE:	1997-06-06
EARNER APPLICATION NUMBER:	60/057,650
EARNER FILING DATE:	1997-09-05

EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 84.9%; Score 1012; DB 4; Length 1049;  
Best Local Similarity 99.3%; Pred. No. 1.2e-263;  
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 46 TCGCGCTGACAGACAGATCTACTCAGCGTGGTCACTCTGTGAACATCATGTACTG 105  
DB 1 TCGCGCTGACAGACAGATCTACTCAGCGTGGTCACTCTGTGAACATCATGTACTG 60  
QY 106 CAGGCTCCCTCAATTTCTGTGACAGCCCATCAAGGACCCCAAGGCTGGAGATG 165  
DB 61 CAGGCTCCCTCAATTTCTGTGACAGCCCATCAAGGACCCCAAGGCTGGAGATG 120  
QY 166 GCGGATCTTGCGCAATGGGGAATGTGACAGACAGACCCCGAGTGAAGACCACTAC 225  
DB 121 GCGGATCTTGCGCAATGGGGAATGTGACAGACAGACCCCGAGTGAAGACCACTAC 180  
QY 226 CCAGCCACCAAGAGTATGATCTCTGACAGAGCTTCTTAATAGGAGCCATGTCTCC 285  
DB 181 CCAGCCACCAAGAGTATGATCTCTGACAGAGCTTCTTAATAGGAGCCATGTCTCC 240  
QY 286 CCGAGGGGGTCTCTGCGCCCGCCAGACAGAGAGTGCAGAGGCTGGTCTGACATC 345  
DB 241 CCGAGGGGGTCTCTGCGCCCGCCAGACAGAGAGTGCAGAGGCTGGTCTGACATC 300  
QY 346 CCGCTTCAATGACCTCAACCGGAGCTGTGAACATGAGGCTTCCGACATGACATCTCG 405  
DB 301 CCGCTTCAATGACCTCAACCGGAGCTGTGAACATGAGGCTTCCGACATGACATCTCG 360  
QY 406 CAACCATGTCTGAGAGCCGCTGACCTTCACTCTCTCTCTCTCTCTCTCTCTCTCT 465  
DB 361 CAACCATGTCTGAGAGCCGCTGACCTTCACTCTCTCTCTCTCTCTCTCTCTCTCT 420  
QY 466 TGTTCGTGCT 525  
DB 421 TGTTCGTGCT 480  
QY 526 ACCTCTGAGGCTGATAGAGGAGTGTGTGAGAGGAACTTGTGAGGCTTGTGTGAG 585  
DB 481 ACCTCTGAGGCTGATAGAGGAGTGTGTGAGAGGAACTTGTGAGGCTTGTGTGAG 540  
QY 586 AGCAGGCTATTTGAGAGGAGTCTGTGTGCTTGAAGTATGATCAAGAGGAGCA 645  
DB 541 AGCAGGCTATTTGAGAGGAGTCTGTGTGCTTGAAGTATGATCAAGAGGAGCA 600  
QY 646 CAGGTGTGCTTCCCTTGTGTGATAGCTGAGAGAGAGAGAGCTTGTGCTTCC--AGCA 703  
DB 601 CAGGTGTGCTTCCCTTGTGTGATAGCTGAGAGAGAGAGAGCTTGTGCTTCC--AGCA 660  
QY 704 TTTCCAAAGTGTGAGTGTGCTGCTGTGCTTCCGAGATCTTTTAAAGTGTATGAGGCTG 763  
DB 661 TTTCCAAAGTGTGAGTGTGCTGCTGTGCTTCCGAGATCTTTTAAAGTGTATGAGGCTG 720  
QY 764 CATTAAGTGGCAAAATCAGAGCAAGAAAGAGATGCCCTTCCCAATTTCTCAATCTT 823  
DB 721 CATTAAGTGGCAAAATCAGAGCAAGAAAGAGATGCCCTTCCCAATTTCTCAATCTT 780  
QY 824 TTAATGCGAAGATCTCAGCTGATGCCAAATGTTCCGATCTGTGAGAGATGCTC 883  
DB 781 TTAATGCGAAGATCTCAGCTGATGCCAAATGTTCCGATCTGTGAGAGATGCTC 840  
QY 884 GAGTCT 943  
DB 841 GAGTCT 900  
QY 944 CTTTATGACATGTTTTTTTAAATGATATTTTAAATGATATTTTAAATGATATTTTAA 1003

DB 901 CTTTATGACATGTTTTTTTAAATGATATTTTAAATGATATTTTAAATGATATTTTAA 960  
QY 1004 ACAAGTGTAGTGGATCTTATTTATTTGACCTTTTCAATTAATTAATTAATTAATTA 1063  
DB 961 ACAAGTGTAGTGGATCTTATTTATTTGACCTTTTCAATTAATTAATTAATTAATTA 1020  
QY 1064 AAAAAAAAAAAAAA 1077  
DB 1021 AAAAAAAAAAAAAA 1034

RESULT 2

US-09-621-976-864  
Sequence 864, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Mline Edwards, J.B.  
APPLICANT: Joberet, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 864  
LENGTH: 483  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURES:  
NAME/KEY: CDS  
LOCATION: 171..482  
US-09-621-976-864

Query Match 39.7%; Score 473; DB 4; Length 483;

Best Local Similarity 100.0%; Pred. No. 3.5e-118; Indels 0; Gaps 0;

Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACCGCTGAGAGCCGCGGAGTGAAGCCGCGGAGCCGAGAGTGGCGCTGACAGCA 60  
DB 11 AAACCGCTGAGAGCCGCGGAGTGAAGCCGCGGAGCCGAGAGTGGCGCTGACAGCA 70  
QY 61 CAGCATCTACTCAGCGTGGGTCACCTCTGTGAACATCACTGACAGACCTTCCCTCAAT 120  
DB 71 CAGCATCTACTCAGCGTGGGTCACCTCTGTGAACATCACTGACAGACCTTCCCTCAAT 130  
QY 121 TTTGTGTGACAGCCATCAAGAGACCAAGCCGCTGAGAGATGAGTGGAGATCTTGGCCAA 180  
DB 131 TTTGTGTGACAGCCATCAAGAGACCAAGCCGCTGAGAGATGAGTGGAGATCTTGGCCAA 190  
QY 181 TGGGGAATCGTGAAG 240  
DB 191 TGGGGAATCGTGAAG 250  
QY 241 TAGCATCTCTGACAGAGCTTCTTAATAGGAGCCATGAGTCTCCCGAGGAGGCTCTGG 300  
DB 251 TAGCATCTCTGACAGAGCTTCTTAATAGGAGCCATGAGTCTCCCGAGGAGGCTCTGG 310  
QY 301 CCGCCGCGAGAGAGAGAGAGTGCAGGCTGGAGTGTGCTGAGTCCCTTCAATGACCT 360  
DB 311 CCGCCGCGAGAGAGAGAGAGTGCAGGCTGGAGTGTGCTGAGTCCCTTCAATGACCT 370  
QY 361 CAACCGGAGAGTGTGAGACATGAGGCTTCCGAGAGAGAGATCTCGGCAACATGCTGTG 420  
DB 371 CAACCGGAGAGTGTGAGACATGAGGCTTCCGAGAGAGAGATCTCGGCAACATGCTGTG 430  
QY 421 GCGGTGACCTTCACT 473  
DB 431 GCGGTGACCTTCACT 483

RESULT 3  
US-09-621-976-3064

Sequence 3064, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 3064  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 93..431  
US-09-621-976-3064

Query Match 32.9%; Score 391.8; DB 4; Length 519;  
Best Local Similarity 98.6%; Pred. No. 3.2e-96;  
Matches 404; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 54 GCAAGACACAGCATCTACTAGCGTGGTGAACCTCTGTGAACATCACTGACAGCTTC 113  
DB 112 GAATACACACAGCATCTACTAGCGTGGTGAACCTCTGTGAACATCACTGACAGCTTC 171  
QY 114 CTTCAATTTTGTGTGAGCCCATCAGGAGCCCAACGCGCTGGAGAGTGGCGGATCT 173  
DB 172 CTTCAATTTTGTGTGAGCCCATCAGGAGCCCAACGCGCTGGAGAGTGGCGGATCT 231  
QY 174 TGGCCATGAGGAAATGTCAGAGACAGACCCCGAGTGAAGACCACTACCCAGCCAC 233  
DB 232 TGGCCATGAGGAAATGTCAGAGACAGACCCCGAGTGAAGACCACTACCCAGCCAC 291  
QY 234 CAAGAGTACATCTCTCGACAGAGCTTTCTCAATGAGGCGCATGCTGCCCGAGGG 293  
DB 292 CAAGAGTACATCTCTCGACAGAGCTTTCTCAATGAGGCGCATGCTGCCCGAGGG 350  
QY 294 GTCCGTGAGCCCGAGCAGAGAGTGCAGGCTGGGTGCTCACTGCCCTTCA 353  
DB 351 GTCCGTGAGCCCGAGCAGAGAGTGCAGGCTGGGTGCTCACTGCCCTTCA 410  
QY 354 ATGACCTCAACCGGAGCTGTGAACATGAGGCTTTTCGAGTGGCATCTCGGCAACATG 413  
DB 411 ATGACCTCAACCGGAGCTGTGAACATGAGGCTTTTCGAGTGGCATCTCGGCAACATG 470  
QY 414 CTGTGAGCGGTGAGCTCATCTCTCTCTTCTGTCATGATGCT 462  
DB 471 CTGTGAGCGGTGAGCTCATCTCTCTCTTCTGTCATGATGCT 519

RESULT 4  
US-09-621-976-17202  
Sequence 17202, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTS and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 17202  
LENGTH: 364  
TYPE: DNA  
ORGANISM: Homo sapiens

Query Match 19.5%; Score 232.8; DB 4; Length 364;  
Best Local Similarity 36.5%; Pred. No. 2.4e-53;  
Matches 122; Conservative 207; Mismatches 4; Indels 1; Gaps 1;

QY 1 AAACCGCTGAGCGCCGAGTGAAGCCCGAGAGCCCGAGTGCCTCTGACAGA 60  
DB 1 AARYSGSKGAPRCGCCGAGAGYGMKSRYSYRRSSCCGSGMGSCKSKRSMRCR 60  
QY 61 CAGCATCTACCAACGCGGTGCTGTGGAACATCACTGACAGCAAGCTCCCTCAAT 120  
DB 61 CMKSWMSMMYMRSMKRYRSTCACTKGGMACMTGSTMIRYASIGMCTSYMARX 120  
QY 121 TTCTGTGACAGCCCATCAGGAGCCAGAGCGCTGTGAGAGATGTGCGCAAT 180  
DB 121 YTCYSKRYMMKYCYKRYRSMCMCMCAGSGMYSRAGSRYSKSGRGMWYMKGCSRA 180  
QY 181 TGGGAAATCGTGAAGACAGACAGCCCGAGTGAAGACCACTACCCAGCCCAAGG 240  
DB 181 TSKGRMMWMMKGRRRATSRYSYGMSSMYGASKRMSMCSASTMSASCMYTMMSAGS 240  
QY 241 TAGCATCTCTGACAGAGCTCTCAATAGGAGCGCATGTGCTCCCGAGAGG-TCCTG 299  
DB 241 YASCAWKSRYRCAKMSCTYSWYMRASMKSKYCAWSRRGSKCYWYSRGSKSCYCWG 300  
QY 300 GCCCGCCGACAGCAGGAGGTGCCAGGCTGG 333  
DB 301 GSCCGCCGACAGCAGGAGGTGCCAGGCTGG 334

RESULT 5  
US-08-821-994-63  
Sequence 63, Application US/08821994A  
Patent No. 6228643  
GENERAL INFORMATION:  
APPLICANT: Greenland, Andrew J  
APPLICANT: Thomas, Ian  
APPLICANT: Jepson, Ian  
TITLE OF INVENTION: Promoters  
FILE REFERENCE: PPD 50108  
CURRENT APPLICATION NUMBER: US/08/821,994A  
CURRENT FILING DATE: 1997-03-22  
EARLIER APPLICATION NUMBER: PCT/GB97/00729  
EARLIER FILING DATE: 1997-03-18  
EARLIER APPLICATION NUMBER: GB 9606062.9  
NUMBER OF SEQ ID NOS: 89  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 63  
LENGTH: 1441  
TYPE: DNA  
ORGANISM: Brassica napus  
US-08-821-994-63

Query Match 4.0%; Score 47.2; DB 3; Length 1441;  
Best Local Similarity 58.6%; Pred. No. 0.0076;  
Matches 82; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 946 TTTTACGACGCTTTTGTGTTAATGATATATTTTATGCTACTTATTTAGTAC 1005  
DB 1289 TGTGTCAATATATATTCATCAAGAAATTTGCTTGTGATGATATTAAG 1348  
QY 1006 AAGTGTAGTGGCATCTATTTATTTGACCTTTCAATTAATAGATTAGTAAAAA 1065  
DB 1349 AAATATATATTAATATATTTTCTCAAAAAAATATATATATATATATATAT 1408  
QY 1066 AAAAAAAAAAAAAAAAAA 1085  
DB 1409 AAAAAAAAAAAAAAAAAA 1428

Sequence 4, Application US/08892770  
Patent No. 5962670  
GENERAL INFORMATION:  
APPLICANT: Walling, Linda L.  
APPLICANT: Pautot, Veronique  
APPLICANT: Gu, Yong-Qiang  
APPLICANT: Chao, Wan Shaw  
TITLE OF INVENTION: Improved Promoters for Enhancing Plant  
TITLE OF INVENTION: Productivity  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/892,770  
FILING DATE: 15-JUL-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Quine, Jonathan A.  
REGISTRATION NUMBER: P-41,261  
REFERENCE/DOCKET NUMBER: 023070-072100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1971 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: -  
LOCATION: 1..1971  
OTHER INFORMATION: /note="tomato acidic leucine  
OTHER INFORMATION: aminopeptidase 2 (Lap2) cDNA"  
US-08-892-770-4  
Query Match  
Best Local Similarity 3.8%; Score 45.8; DB 2; Length 1971;  
Matches 89; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 925 AGGTGAAAAAATCCAGACTTTTGAACGCTGTTTGAATGTAATTTTATT 984  
DB 1806 AGATGAATAAATGCGTAATTAATTTGCACTTTCAATGAATTTATTATT 1865  
QY 985 GGCTACTTTTGTGTTAGACCAAGTGTAAGTGCATTCTATTATGTGACCTTTCAAT 1044  
DB 1866 CCGCTCTATTATGCTATGCTGATGAGATTATATTCCTCTAGATGTAACCTTGAG 1925  
QY 1045 AATAGATTAACTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1085  
DB 1926 ATGAAGTGTATAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1966  
RESULT 7  
US-09-904-615-66  
Sequence 66, Application US/09904615  
Patent No. 656325  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 49 Human Secreted Proteins  
FILE REFERENCE: P2032P1  
CURRENT APPLICATION NUMBER: US/09/904,615

CURRENT FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 09/511,554  
PRIOR FILING DATE: 2000-02-23  
PRIOR APPLICATION NUMBER: 60/097,917  
PRIOR FILING DATE: 1998-08-25  
PRIOR APPLICATION NUMBER: 60/098,634  
PRIOR FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 170  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 66  
LENGTH: 664  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (31)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (63)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-904-615-66  
Query Match  
Best Local Similarity 3.8%; Score 45.6; DB 4; Length 664;  
Matches 87; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 930 GAAAAAATTCAGACTTTTGAACGCTGTTTGAATGTAATTTTATTGGCTA 989  
DB 493 GAAAAAATTTAATATTGTAAGGTTTTCCTTTTATTTTATTTTGGTATA 552  
QY 990 CTTATGTTTGAACAAGTGTAAGTGCATTCTATTATTTGTAACCTTTCAATAATA 1049  
DB 553 TTGCTGTATCTAATTAACTTCAGAAATTAACGTTATATGGAATTTTAAAAA 612  
QY 1050 GATTAAAGTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1085  
DB 613 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 648  
RESULT 8  
US-09-293-322C-8  
Sequence 8, Application US/09293322C  
Patent No. 6232110  
GENERAL INFORMATION:  
APPLICANT: Pallas, David C  
APPLICANT: Du, Xianxing  
TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methyltransferase,  
Patent No. 6232110  
TITLE OF INVENTION: Recombinant DNA Molecules and Methods  
FILE REFERENCE: 105-97  
CURRENT APPLICATION NUMBER: US/09/293,322C  
CURRENT FILING DATE: 1999-04-16  
PRIOR APPLICATION NUMBER: US 60/082,202  
PRIOR FILING DATE: 1998-04-17  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 8  
LENGTH: 2409  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(2409)  
OTHER INFORMATION: N is A, T, G or C.  
US-09-293-322C-8  
Query Match  
Best Local Similarity 3.8%; Score 45.4; DB 3; Length 2409;  
Matches 67; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
QY 983 TTGGCTACTTTATTTGTTTGAACAAGTGTAAGTGCATTCTATTATTTGTAACCTTTTCA 1042  
DB 2294 TTGCTCTTTGTTATTTATGATCTGTGTTAAAGAAAAATATATCTCCACCTTTAAA 2353

Query 1043 ATAAATGATTAAAGTAAAAA 1085  
Db 2354 AAAAAAAAAAAAAAAAAA 2396

## RESULT 9

US-09-839-497A-8  
Sequence 8, Application US/09839497A  
Patent No. 6528295  
GENERAL INFORMATION:  
APPLICANT: Pallas, David C.  
APPLICANT: Du, Xianxing  
TITLE OF INVENTION: Coding Sequence for Protein Phosphatase Methyltransferase,  
Patent No. 6528295  
TITLE OF INVENTION: Recombinant DNA Molecules and Methods  
FILE REFERENCES: Docket No. 6528295 105-97A  
CURRENT APPLICATION NUMBER: US/09/839,497A  
CURRENT FILING DATE: 2001-04-20  
PRIOR APPLICATION NUMBER: 60/082,202  
PRIOR FILING DATE: 1998-04-17  
PRIOR APPLICATION NUMBER: 09/293,322  
PRIOR FILING DATE: 1999-04-16  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 8  
LENGTH: 2409  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(2409)  
OTHER INFORMATION: N is A, T, G or C.  
US-09-839-497A-8

Query Match 3.8%; Score 45.4; DB 4; Length 2409;  
Best Local Similarity 65.0%; Pred. No. 0.031;  
Matches 6; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Query 983 TTGGCTACTTATTTAGTGAAGTGTAGTGGATTCTATTATTGTGACCTTTTCA 1042  
Db 2294 TTTGTTCTTTGTTATTTATGATCTGTTTAAAGAAATATATCTCCCACTTTAA 2353  
Query 1043 ATAAATGATTAAAGTAAAAA 1085  
Db 2354 AAAAAAAAAAAAAAAAAA 2396

## RESULT 10

US-09-889-463A-9  
Sequence 9, Application US/09889463A  
Patent No. 6680185  
GENERAL INFORMATION:  
APPLICANT: Cahoon, Rebecca E.  
APPLICANT: Falco, Savetio C.  
APPLICANT: Kinney, Anthony J.  
APPLICANT: Miao, Guo-Hua  
TITLE OF INVENTION: Plant Polyphenol Oxidase Homologs  
FILE REFERENCES: BB1330  
CURRENT APPLICATION NUMBER: US/09/889,463A  
CURRENT FILING DATE: 2001-07-16  
PRIOR APPLICATION NUMBER: 60/119,590  
PRIOR FILING DATE: 1999-02-10  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 9  
LENGTH: 2485  
TYPE: DNA  
ORGANISM: Glycine max  
US-09-889-463A-9

Query Match 3.8%; Score 45.2; DB 4; Length 2485;  
Best Local Similarity 55.8%; Pred. No. 0.036;

Matches 86; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Query 932 AAAATTCAGACTTTTGGACTGTTTGTGTAAGTATTTATGGCTACT 991  
Db 1947 ATATPAAGTGTCTTGATTTGTTGTTTATATGAAAAATAGTTTGTAT 2006  
Query 992 TTATGTTAGACAGTGTAGTGCATTCTATTATGACCTTTCAATPAATGA 1051  
Db 2007 GTATTTTCATGAGAACATPAAGATGTCCTATATGAGTGAAGAAAAA 2066  
Query 1052 TTTAAGTAAAAA 1085  
Db 2067 AAAAAAAAAA 2100

## RESULT 11

US-07-867-106-4  
Sequence 4, Application US/07867106  
Patent No. 5389526  
GENERAL INFORMATION:  
APPLICANT: Slade, Martin B  
APPLICANT: Chang, Andy C M  
APPLICANT: Williams, Keith L  
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular  
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris  
STREET: One Liberty Place 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/867,106  
FILING DATE: 19920625  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PJ 7187  
APPLICATION NUMBER: PCT/AU90/00530  
FILING DATE: 02-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Feeney, Joanne Longo  
REGISTRATION NUMBER: 35,134  
REFERENCE/DOCKET NUMBER: RICE-0002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3439  
TELEFAX: 215-568-3100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3138 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
ANTI-SENSE: NO  
US-07-867-106-4

Query Match 3.8%; Score 45.2; DB 1; Length 3138;  
Best Local Similarity 54.9%; Pred. No. 0.041;  
Matches 89; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

Query 924 CAGGTGAAAAAATTCAGACTTTTGTAGCACTGTTTGTGTAAGTATTTAT 983  
Db 1912 CAAATTAATAAATAAATGGAATCAAAATTTTATTTTATTTTATTTT 1971  
Query 984 TGGCTACTTATTTTGAAGCAAGTGTAGTGCATTCTATTATGACCTTTCA 1043  
Db 1972 TTTTATTTTATTTTTCAGATTCATPAAGATTTTATTTTATTTTCTATTTTC 2031



**Tue Aug 24 09:48:34 2004**

us-10-001-885-31.rml

Page 8

OTHER INFORMATION: 470 bp variable  
OTHER INFORMATION: region where R is A or G  
US-08-021-608D-9

Query Match	3.8%	Score 44.8;	DB 1;	length 2381;
Best Local Similarity	59.4%;	Pred. No. 0.045;		
Matches 76;	Conservative 0;	Mismatches 52;	Indels 0;	Gaps 0;

Accession	Sequence	Position
QY	1018 CATCATATTTATTTGACCTTTTCAATTAATAGATTTAAGTAAAAA	10778
Dp	2236 TTGTGATTTAAATATACCTTTTCAATGTAACCTTACCTTCAAAATCCGTGTTGTG	22955
QY	938 TTTTGTGTTTAAATGTAATTTTAAATGACATCTTAAATGAGCAAGGTAAG	10778
Dp	2236 TTGTGATTTAAATATACCTTTTCAATGTAACCTTACCTTCAAAATCCGTGTTGTG	22955
Dp	2296 CTTTCAATTAATATGATATGAAACCTCAAAAAA	23555

QY	1078	AAAAA	1085
Db	2356	AAAAAAA	2363

RESULT 15  
US-08-726-160-9

GENERAL INFORMATION:  
APPLICANT: LEVENS

APPLICANT: LEVENS, DAVID L., DUNCAN,  
APPLICANT: ROBERT C., AND AVIGAN, MARK I.

TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR  
NUMBER OF SEQUENCES: 24

ADDRESSEE: MORGAN & FINNEGAN

CITY: NEW YORK  
STATE: NEW YORK

```

?      COMPUTER READABLE FORM:
?      MEDIUM TYPE: FLOPPY DISK
?      COMPUTER: IBM PC COMPATIBLE
?      OPERATING SYSTEM: PC-DOS/MS-DOS
?      SOFTWARE: WORDPERFECT 5.1

```

CONSENT AFFIDAVIT DATA.  
 APPLICATION NUMBER: US/08/726,160  
 FILING DATE: 04-OCT-1996

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/021,608
; FILING DATE: 08/08/2008

```

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4063US1

TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2381

```

STRANDEDNESS: Double  
TOPOLOGY: Unknown

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

ORGANISM: Human  
CELL LINE: HL60  
FEATURE:

```

OTHER INFORMATION: 470 bp variable
OTHER INFORMATION: region where R is A or G.

```

US-08-726-160-9

Query Match	3.8%	Score 44.8	DB 1	Length 2381
Best Local Similarity	59.4%	Pred. No. 0.045		
Matches	76	Conservative	0	Mismatches 52
				Indels 0
				Gaps 0

[illegible]

```
QY      1078 AAAAAAAAAA 1085
          |||||
Db      2356 AAAAAAAAAA 2363
```

```
Search completed: August 22, 2004, 03:31:26
Job time : 122 secs
```



```

RESULT 1
US-10-001-885-31
; Sequence 31: Application US/10001885
; Publication No. US20040058319A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Cafierkey, Robert
; APPLICANT: Liu, Chengshua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; FILE REFERENCE: DEX-0279
; CURRENT APPLICATION NUMBER: US/10/001,885
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,061
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/253,257
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 1192
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-001-885-31

Query Match      100.0%; Score 1192; DB 13; Length 1192;
Best Local Similarity 100.0%; Pred. No. 6,9e-309;
Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AAACCGCCTGACCCGCGCGGAGTGGACGCGCGCGAGAGTCCGCGCTGCAGACA 60
Db      1 AAACCGCCTGAGACCGCGCGGAGTGGACGCGCGCGAGTCCGCGCTGCAGACA 60
QY      61 CAGCATCTATCAGCGGTGAGTCACTCTGTGAACATCAGCATGGAAGCTCCCTCAAT 120

```

Db	61	CAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTAGTCGACGCTCCCTCAAT	120
Qy	121	TTCTGGTGCAGCCCATCAGGGAACCAAGAGCTTGGAGGATAGTGTGGATCTTGGCCAA	180
Db	121	TTCTGGTGCAGCCCATCAGGGAACCAAGAGCTTGGAGGATAGTGTGGATCTTGGCCAA	180
Qy	181	TGGGAAATCGTGCAGAGCAGACACCCCGAGTGAACCACTACCCAGCCCAAGAG	240
Db	181	TGGGAAATCGTGCAGAGCAGACACCCCGAGTGAACCACTACCCAGCCCAAGAG	240
Qy	241	TAGCATTTCCCTGACAGAGCTTCTTCAATAGGGGCAATGGTGTCCCGAGGGGGCTCTGG	300
Db	241	TAGCATTTCCCTGACAGAGCTTCTTCAATAGGGGCAATGGTGTCCCGAGGGGGCTCTGG	300
Qy	301	CCCCCGCAGCAGCAGGAGGTGCGAGCTGGTGTCTAGTCCCTCCCTCAATGACCT	360
Db	301	CCCCCGCAGCAGCAGGAGGTGCGAGCTGGTGTCTAGTCCCTCCCTCAATGACCT	360
Qy	361	CAACCCGACAGTGTGAACATGGGCTTTCCGAGTGGCATCTTGGCAACATGCTGTGA	420
Db	361	CAACCCGACAGTGTGAACATGGGCTTTCCGAGTGGCATCTTGGCAACATGCTGTGA	420
Qy	421	GGCGGTGACCTGCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	480
Db	421	GGCGGTGACCTGCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	480
Qy	481	CTGTGTGGCTGTCTACTGTGTCTCACTGTGTCTCACTGTGTGTGTGTGTGTGTGTGT	540
Db	481	CTGTGTGGCTGTCTACTGTGTCTCACTGTGTCTCACTGTGTGTGTGTGTGTGTGTGT	540
Qy	541	TAGGGGTGGGT	600
Db	541	TAGGGGTGGGT	600
Qy	601	AGGGGATCTGT	660
Db	601	AGGGGATCTGT	660
Qy	661	CTTTGT	720
Db	661	CTTTGT	720
Qy	721	GTCTGT	780
Db	721	GTCTGT	780
Qy	781	TGAGAGCAGAAAGCGATGCTTCCCAATCTCTCAATCTCTTATGCGGAGAGATCT	840
Db	781	TGAGAGCAGAAAGCGATGCTTCCCAATCTCTCAATCTCTTATGCGGAGAGATCT	840
Qy	841	CAGGTGTGACCAATGTTCGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	900
Db	841	CAGGTGTGACCAATGTTCGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	900
Qy	901	GGGAGCAGACTTGGGCTTAAAGGAGGTGTGAAAAATTCAGACTTTTAAAGCTGTTT	960
Db	901	GGGAGCAGACTTGGGCTTAAAGGAGGTGTGAAAAATTCAGACTTTTAAAGCTGTTT	960
Qy	961	TTGTTTAAATGT	1020
Db	961	TTGTTTAAATGT	1020
Qy	1021	TCTATTATGTGACCTTTTCAATTAATAGATTAAATTAATTAATTAATTAATTAATTA	1080
Db	1021	TCTATTATGTGACCTTTTCAATTAATAGATTAAATTAATTAATTAATTAATTAATTA	1080
Qy	1081	AAAAAGTGTGGGGTTTACCCGGGGCAAAAGGGGGCCCCGGGGGGGAATGTGTGTCTG	1140
Db	1081	AAAAAGTGTGGGGTTTACCCGGGGCAAAAGGGGGCCCCGGGGGGGAATGTGTGTCTG	1140
Qy	1141	CGGCCCCAATTCCTCCCAATTTTGA CAAATGAGAGGACACCGGAAACC 1192	
Db	1141	CGGCCCCAATTCCTCCCAATTTTGA CAAATGAGAGGACACCGGAAACC 1192	

Qy	15	CGCGGGAGTGAAGGAGCGCGCGGAGGCGCGAGTGGCGGCTGCAGACAGATCTACTAG	74
Db	1	CGCGGGAGTGAAGGAGCGCGCGGAGGCGCGAGTGGCGGCTGCAGACAGATCTACTAG	60
Qy	75	CGTGGTGCACCTCTGTGAACATCACTGACTGCAAGCTCTCCATTTCTGTGTGAGCCC	134
Db	61	CGTGGTGCACCTCTGTGAACATCACTGACTGCAAGCTCTCCATTTCTGTGTGAGCCC	120
Qy	135	ATCAGGAGCCACAGCGCCCTGTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	194
Db	121	ATCAGGAGCCACAGCGCCCTGTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	180
Qy	195	AGGAGCAGACCCCGAGTGAAGGACACTACCCAGCCAGCAGAGGTAGATTCCTGAC	254
Db	181	AGGAGCAGACCCCGAGTGAAGGACACTACCCAGCCAGCAGAGGTAGATTCCTGAC	240
Qy	255	AGAGCTTCTCAATAGGGGCGCATGTGTCTCCCAAGGGGTCTGTGGCCCCCGCAGAGC	314
Db	241	AGAGCTTCTCAATAGGGGCGCATGTGTCTCCCAAGGGGTCTGTGGCCCCCGCAGAGC	300
Qy	315	AGGAGGAGCAGGCTGGGT	374
Db	301	AGGAGGAGCAGGCTGGGT	360
Qy	375	TGAACATGGGCTTTCGAGT	434
Db	361	TGAACATGGGCTTTCGAGT	420
Qy	435	TCCGTCT	494
Db	421	TCCGTCT	480
Qy	495	TCTACTGT	554
Db	481	TCTACTGT	540
Qy	555	TTGAGAGGAGCTTGT	614
Db	541	TTGAGAGGAGCTTGT	600
Qy	615	TGCTTGAAGGTATGATCAGAGGAGCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	674

Query Match 89.0%; Score 1061.4; DB 16; Length 1108;  
Best Local Similarity 99.6%; Pred. No. 7.2e-274;  
Matches 1059; Conservative 4; Mismatches 0; Indels 0; Gaps 0;



FILE REFERENCE: P2002P2  
CURRENT APPLICATION NUMBER: US/09/809,391  
CURRENT FILING DATE: 2001-03-16  
Prior application data removed - consult PAM or file wrapper  
NUMBER OF SEQ ID NOS: 761  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 58  
LENGTH: 1049  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-809-391-58

Query Match 84.9%; Score 1012; DB 10; Length 1049;  
Best Local Similarity 99.3%; Pred. No. 1.3e-260;  
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 46 TCGGCGCTGAGAGACAGCATCTACTACAGGCTGGGTCACTCTGTAAACATCACTGACTG 105  
DB 1 TCGGCGCTGAGAGACAGCATCTACTACAGGCTGGGTCACTCTGTAAACATCACTGACTG 60  
QY 106 CAACCTCTCCCTCATTTCTGGTGCAGCCCATCAGGACCCACAGCGCTGGAGATGGT 165  
DB 61 CAACCTCTCCCTCATTTCTGGTGCAGCCCATCAGGACCCACAGCGCTGGAGATGGT 120  
QY 166 GCGGATCTGGCCCATGGGAAATCTGTGCAGACAGACAGACCCCGAGTGAAGCACTAC 225  
DB 121 GCGGATCTGGCCCATGGGAAATCTGTGCAGACAGACAGACCCCGAGTGAAGCACTAC 180  
QY 226 CCAGCCACAGAGAGGTGACATCTCTGCAGAGCTTTCAATAGGGGCCATGGTCTCC 285  
DB 181 CCAGCCACAGAGAGGTGACATCTCTGCAGAGCTTTCAATAGGGGCCATGGTCTCC 240  
QY 286 CCCAGGGGGTCTGTCGCCCGCCAGCAGCAGAGGAGGTGCCAGGCTGGTCTGCTAGTC 345  
DB 241 CCCAGGGGGTCTGTCGCCCGCCAGCAGCAGAGGAGGTGCCAGGCTGGTCTGCTAGTC 300  
QY 346 CCCCTTAATGACCTCAACCGGAGCTGTGAAATGAGGCTTCCGAGTGGATCTCGG 405  
DB 301 CCCCTTAATGACCTCAACCGGAGCTGTGAAATGAGGCTTCCGAGTGGATCTCGG 360  
QY 406 CAACCATGCTGTGAGCGCGGTGACCTTCATCTGCTCTCTTCTCTCTCATGATGCTGG 465  
DB 361 CAACCATGCTGTGAGCGCGGTGACCTTCATCTGCTCTCTTCTCTCTCATGATGCTGG 420  
QY 466 TGTTCGTGGCTCTCTCTGCTGTGACCTGTGACCTGTGACCTGTGACCTGTGACCTGTG 525  
DB 421 TGTTCGTGGCTCTCTCTGCTGTGACCTGTGACCTGTGACCTGTGACCTGTGACCTGTG 480  
QY 526 ACCTCTAGAGGGCTGATAGGGGTGGTGTGTTGAGAGGGAATTGCTGGGCTGGTGGAG 585  
DB 481 ACCTCTAGAGGGCTGATAGGGGTGGTGTGTTGAGAGGGAATTGCTGGGCTGGTGGAG 540  
QY 586 AGCAGGCAATTTTGGAGGGGATCTGGTGTGCTTTAAGGATGATCAAGAGGGGACCA 645  
DB 541 AGCAGGCAATTTTGGAGGGGATCTGGTGTGCTTTAAGGATGATCAAGAGGGGACCA 600  
QY 646 CAGGTGTGTTTCCCTTTGTTAAGGTTGAGGAGAGGGGAGAGCTTAAGTCC--AGCA 703  
DB 601 CAGGTGTGTTTCCCTTTGTTAAGGTTGAGGAGAGGGGAGAGCTTAAGTCC--AGCA 660  
QY 704 TTTCCAAAGTGTGGGTGGGTCCGTGTCCGAGATCTTTTGTAGGTGTAGGGGCTGTG 763  
DB 661 TTTCCAAAGTGTGGGTGGGTCCGTGTCCGAGATCTTTTGTAGGTGTAGGGGCTGTG 720  
QY 764 CATTAGTGGCAAAATGAGAGAGAAAGAGATCCCTTCCCATTTCTTCATCTCT 823  
DB 721 CATTAGTGGCAAAATGAGAGAGAAAGAGATCCCTTCCCATTTCTTCATCTCT 780  
QY 824 TTATGCGAGAGATCTCAGCTGTGATGCCAAACATGTTCCGATGCTGTGAGAGCATGCC 883  
DB 781 TTATGCGAGAGATCTCAGCTGTGATGCCAAACATGTTCCGATGCTGTGAGAGCATGCC 840

DB 841 GACGCTCTCTGCTAGAGAGAGAGACTTGGGCTTAGGGCAGGTGGAAAAATTCAGCA 900  
QY 944 CTTTTTACGACTGTTTTTGTGTTAATGATATTTTATGCTACTTTATGTTAGG 1003  
DB 901 CTTTTTACGACTGTTTTTGTGTTAATGATATTTTATGCTACTTTATGTTAGG 960  
QY 1004 ACAAGTGTAGTGCATCTATTATTGTGACCTTTCAATTAATGATTTAAGTAAAA 1063  
DB 961 ACAAGTGTAGTGCATCTATTATTGTGACCTTTCAATTAATGATTTAAGTAAAA 1020

RESULT 5  
US-09-882-171-58  
Sequence 58, Application US/09882171

Publication No. US20030175858A1  
GENERAL INFORMATION:

APPLICANT: Ruben et al.  
TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P2  
CURRENT FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 09/809,391

PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 09/149,476  
PRIOR FILING DATE: 1998-09-08

PRIOR APPLICATION NUMBER: PCT/US98/04493  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: 60/040,162

PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,333  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/038,621  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,626

PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,334  
PRIOR FILING DATE: 1997-03-07

PRIOR APPLICATION NUMBER: 60/040,336  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,163

PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/047,600  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,615  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,597

PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,502  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,633  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,583

PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,617  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,618  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,503

PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,592  
PRIOR FILING DATE: 1997-05-23

PRIOR APPLICATION NUMBER: 60/047,581  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,584

PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,500  
PRIOR FILING DATE: 1997-05-23

1 PRIOR FILING DATE: 1997-05-23  
2 PRIOR APPLICATION NUMBER: 60/047,492  
3 PRIOR FILING DATE: 1997-05-23  
4 PRIOR APPLICATION NUMBER: 60/047,598  
5 PRIOR FILING DATE: 1997-05-23  
6 PRIOR APPLICATION NUMBER: 60/047,613  
7 PRIOR FILING DATE: 1997-05-23  
8 PRIOR APPLICATION NUMBER: 60/047,582  
9 PRIOR FILING DATE: 1997-05-23  
10 PRIOR APPLICATION NUMBER: 60/047,596  
11 PRIOR FILING DATE: 1997-05-23  
12 PRIOR APPLICATION NUMBER: 60/047,612  
13 PRIOR FILING DATE: 1997-05-23  
14 PRIOR APPLICATION NUMBER: 60/047,632  
15 PRIOR FILING DATE: 1997-05-23  
16 PRIOR APPLICATION NUMBER: 60/047,601  
17 PRIOR FILING DATE: 1997-05-23  
18 PRIOR APPLICATION NUMBER: 60/043,580  
19 PRIOR FILING DATE: 1997-04-11  
20 PRIOR APPLICATION NUMBER: 60/043,568  
21 PRIOR FILING DATE: 1997-04-11  
22 PRIOR APPLICATION NUMBER: 60/043,314  
23 PRIOR FILING DATE: 1997-04-11  
24 PRIOR APPLICATION NUMBER: 60/043,569  
25 PRIOR FILING DATE: 1997-04-11  
26 PRIOR APPLICATION NUMBER: 60/043,311  
27 PRIOR FILING DATE: 1997-04-11  
28 PRIOR APPLICATION NUMBER: 60/043,671  
29 PRIOR FILING DATE: 1997-04-11  
30 PRIOR APPLICATION NUMBER: 60/043,674  
31 PRIOR FILING DATE: 1997-04-11  
32 PRIOR APPLICATION NUMBER: 60/043,669  
33 PRIOR FILING DATE: 1997-04-11  
34 PRIOR APPLICATION NUMBER: 60/043,312  
35 PRIOR FILING DATE: 1997-04-11  
36 PRIOR APPLICATION NUMBER: 60/043,313  
37 PRIOR FILING DATE: 1997-04-11  
38 PRIOR APPLICATION NUMBER: 60/043,672  
39 PRIOR FILING DATE: 1997-04-11  
40 PRIOR APPLICATION NUMBER: 60/043,315  
41 PRIOR FILING DATE: 1997-04-11  
42 PRIOR APPLICATION NUMBER: 60/048,974  
43 PRIOR FILING DATE: 1997-06-06  
44 PRIOR APPLICATION NUMBER: 60/056,886  
45 PRIOR FILING DATE: 1997-08-22  
46 PRIOR APPLICATION NUMBER: 60/056,877  
47 PRIOR FILING DATE: 1997-08-22  
48 PRIOR APPLICATION NUMBER: 60/056,889  
49 PRIOR FILING DATE: 1997-08-22  
50 PRIOR APPLICATION NUMBER: 60/056,893  
51 PRIOR FILING DATE: 1997-08-22  
52 PRIOR APPLICATION NUMBER: 60/056,630  
53 PRIOR FILING DATE: 1997-08-22  
54 PRIOR APPLICATION NUMBER: 60/056,878  
55 PRIOR FILING DATE: 1997-08-22  
56 PRIOR APPLICATION NUMBER: 60/056,662  
57 PRIOR FILING DATE: 1997-08-22  
58 PRIOR APPLICATION NUMBER: 60/056,872  
59 PRIOR FILING DATE: 1997-08-22  
60 PRIOR APPLICATION NUMBER: 60/056,882  
61 PRIOR FILING DATE: 1997-08-22  
62 PRIOR APPLICATION NUMBER: 60/056,637  
63 PRIOR FILING DATE: 1997-08-22  
64 PRIOR APPLICATION NUMBER: 60/056,903  
65 PRIOR FILING DATE: 1997-08-22  
66 PRIOR APPLICATION NUMBER: 60/056,888  
67 PRIOR FILING DATE: 1997-08-22  
68 PRIOR APPLICATION NUMBER: 60/056,879  
69 PRIOR FILING DATE: 1997-08-22  
70 PRIOR APPLICATION NUMBER: 60/056,880  
71 PRIOR FILING DATE: 1997-08-22  
72 PRIOR APPLICATION NUMBER: 60/056,894  
73 PRIOR FILING DATE: 1997-08-22

74 PRIOR APPLICATION NUMBER: 60/056,911  
75 PRIOR FILING DATE: 1997-08-22  
76 PRIOR APPLICATION NUMBER: 60/056,636  
77 PRIOR FILING DATE: 1997-08-22  
78 PRIOR APPLICATION NUMBER: 60/056,874  
79 PRIOR FILING DATE: 1997-08-22  
80 PRIOR APPLICATION NUMBER: 60/056,910  
81 PRIOR FILING DATE: 1997-08-22  
82 PRIOR APPLICATION NUMBER: 60/056,864  
83 PRIOR FILING DATE: 1997-08-22  
84 PRIOR APPLICATION NUMBER: 60/056,631  
85 PRIOR FILING DATE: 1997-08-22  
86 PRIOR APPLICATION NUMBER: 60/056,845  
87 PRIOR FILING DATE: 1997-08-22  
88 PRIOR APPLICATION NUMBER: 60/056,892  
89 PRIOR FILING DATE: 1997-08-22  
90 PRIOR APPLICATION NUMBER: 60/057,761  
91 PRIOR FILING DATE: 1997-08-22  
92 PRIOR APPLICATION NUMBER: 60/047,595  
93 PRIOR FILING DATE: 1997-05-23  
94 PRIOR APPLICATION NUMBER: 60/047,599  
95 PRIOR FILING DATE: 1997-05-23  
96 PRIOR APPLICATION NUMBER: 60/047,588  
97 PRIOR FILING DATE: 1997-05-23  
98 PRIOR APPLICATION NUMBER: 60/047,585  
99 PRIOR FILING DATE: 1997-05-23  
100 PRIOR APPLICATION NUMBER: 60/047,586  
101 PRIOR FILING DATE: 1997-05-23  
102 PRIOR APPLICATION NUMBER: 60/047,590  
103 PRIOR FILING DATE: 1997-05-23  
104 PRIOR APPLICATION NUMBER: 60/047,594  
105 PRIOR FILING DATE: 1997-05-23  
106 PRIOR APPLICATION NUMBER: 60/047,589  
107 PRIOR FILING DATE: 1997-05-23  
108 PRIOR APPLICATION NUMBER: 60/047,593  
109 PRIOR FILING DATE: 1997-05-23  
110 PRIOR APPLICATION NUMBER: 60/047,614  
111 PRIOR FILING DATE: 1997-05-23  
112 PRIOR APPLICATION NUMBER: 60/043,578  
113 PRIOR FILING DATE: 1997-04-11  
114 PRIOR APPLICATION NUMBER: 60/043,576  
115 PRIOR FILING DATE: 1997-04-11  
116 PRIOR APPLICATION NUMBER: 60/047,501  
117 PRIOR FILING DATE: 1997-05-23  
118 PRIOR APPLICATION NUMBER: 60/043,670  
119 PRIOR FILING DATE: 1997-04-11  
120 PRIOR APPLICATION NUMBER: 60/056,632  
121 PRIOR FILING DATE: 1997-08-22  
122 PRIOR APPLICATION NUMBER: 60/056,664  
123 PRIOR FILING DATE: 1997-08-22  
124 PRIOR APPLICATION NUMBER: 60/056,876  
125 PRIOR FILING DATE: 1997-08-22  
126 PRIOR APPLICATION NUMBER: 60/056,881  
127 PRIOR FILING DATE: 1997-08-22  
128 PRIOR APPLICATION NUMBER: 60/056,909  
129 PRIOR FILING DATE: 1997-08-22  
130 PRIOR APPLICATION NUMBER: 60/056,875  
131 PRIOR FILING DATE: 1997-08-22  
132 PRIOR APPLICATION NUMBER: 60/056,862  
133 PRIOR FILING DATE: 1997-08-22  
134 PRIOR APPLICATION NUMBER: 60/056,887  
135 PRIOR FILING DATE: 1997-08-22  
136 PRIOR APPLICATION NUMBER: 60/056,908  
137 PRIOR FILING DATE: 1997-08-22  
138 PRIOR APPLICATION NUMBER: 60/048,964  
139 PRIOR FILING DATE: 1997-06-06  
140 PRIOR APPLICATION NUMBER: 60/057,650  
141 PRIOR FILING DATE: 1997-09-05  
142 PRIOR APPLICATION NUMBER: 60/056,884  
143 PRIOR FILING DATE: 1997-08-22  
144 PRIOR APPLICATION NUMBER: 60/057,669  
145 PRIOR FILING DATE: 1997-09-05

Query Match 84.9%; Score 1012; DB 10; Length 1049;  
Best Local Similarity 99.3%; Pred. No. 1,3e-260;  
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

```
QY 46 TCGGCGCTGACAGACAGCATCTAAGCGTGGGTCACCTGTGTGAACATCACTGACTG 105
Db 1 TCGGCGCTGACAGACAGCATCTAAGCGTGGGTCACCTGTGTGAACATCACTGACTG 60
QY 106 CAAGCCTCCCTCAATTTCTGTGTGAGGCCCAATCAAGGGGCCCAAGCGCTGGGAGATGGT 165
Db 61 CAAGCCTCCCTCAATTTCTGTGTGAGGCCCAATCAAGGGGCCCAAGCGCTGGGAGATGGT 120
QY 166 GCGGATCTGGCCAAATGAGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 225
Db 121 GCGGATCTGGCCAAATGAGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 226 CCAAGCCCAAGAGAGATGAGATCTCTGACAGAGCTTTCAATAGAGGAGCATGTGCTCC 285
Db 181 CCAAGCCCAAGAGAGATGAGATCTCTGACAGAGCTTTCAATAGAGGAGCATGTGCTCC 240
QY 286 CCGAGGGGGGTCCTGGCCCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
Db 241 CCGAGGGGGGTCCTGGCCCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 346 CCGCTTCAATGACCTCAACCGGAGAGCTGTGAACATGAGGCTTTCCGAGATGAGCTCCG 405
Db 301 CCGCTTCAATGACCTCAACCGGAGAGCTGTGAACATGAGGCTTTCCGAGATGAGCTCCG 360
QY 406 CAACCAATGCTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
Db 361 CAACCAATGCTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 466 TGTTCGTGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
Db 421 TGTTCGTGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 526 ACCTCTGAGGAGCTATAGAGGAGTGGGTTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 585
Db 481 ACCTCTGAGGAGCTATAGAGGAGTGGGTTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 586 AGCAGGATATTTGAGAGGAGATCTGTGTGAGGCTGTGAAGATGATCAAGAGAGAGAG 645
Db 541 AGCAGGATATTTGAGAGGAGATCTGTGTGAGGCTGTGAAGATGATCAAGAGAGAGAG 600
QY 646 CAGGTGTGTGCTCCCTTTGTGTGAAGCTGTGAAGAGAGAGAGAGAGAGAGAGAGAG 703
Db 601 CAGGTGTGTGCTCCCTTTGTGTGAAGCTGTGAAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 704 TTTTCAAAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 763
Db 661 TTTTCAAAAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAG 720
QY 764 CATTAGTGGACAAAATCAGAGCAAGAAAGCATGCGCTTCCCAATCTCTCAATCCTT 823
Db 721 CATTAGTGGACAAAATCAGAGCAAGAAAGCATGCGCTTCCCAATCTCTCAATCCTT 780
QY 824 TTTATGCCGAAATATCTCAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 883
Db 781 TTTATGCCGAAATATCTCAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 884 GAGCTCTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 943
Db 841 GAGCTCTCTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 944 CTTTTTATGAGAGCTGTTTTTTTATATGATATTTTATATGAGAGAGAGAGAGAGAG 1003
Db 901 CTTTTTATGAGAGCTGTTTTTTTATATGATATTTTATATGAGAGAGAGAGAGAGAG 960
QY 1004 ACAAGTGTGTGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1063
Db 961 ACAAGTGTGTGAGAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1020
```

Db 1021 AAAAAAAAAAAAAA 1014

```
RESULT 6
US-10-164-861-58
; Sequence 58, Application US/10164861
; Publication No. US20030225248A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/10/164, 861
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US/09/149, 476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; NUMBER OF SEQ ID NOS: 757
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 1049
; TYPR: DNA
; ORGANISM: Homo sapiens
US-10-164-861-58
```

Query Match 84.9%; Score 1012; DB 13; Length 1049;  
Best Local Similarity 99.3%; Pred. No. 1,3e-260;  
Matches 1027; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

```
QY 46 TCGGCGCTGACAGACAGCATCTAAGCGTGGGTCACCTGTGTGAACATCACTGACTG 105
Db 1 TCGGCGCTGACAGACAGCATCTAAGCGTGGGTCACCTGTGTGAACATCACTGACTG 60
QY 106 CAAGCCTCCCTCAATTTCTGTGTGAGGCCCAATCAAGGGGCCCAAGCGCTGGGAGATGGT 165
Db 61 CAAGCCTCCCTCAATTTCTGTGTGAGGCCCAATCAAGGGGCCCAAGCGCTGGGAGATGGT 120
QY 166 GCGGATCTGGCCAAATGAGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 225
Db 121 GCGGATCTGGCCAAATGAGGAAATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 226 CCAAGCCCAAGAGAGATGAGATCTCTGACAGAGCTTTCAATAGAGGAGCATGTGCTCC 285
Db 181 CCAAGCCCAAGAGAGATGAGATCTCTGACAGAGCTTTCAATAGAGGAGCATGTGCTCC 240
QY 286 CCGAGGGGGGTCCTGGCCCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
Db 241 CCGAGGGGGGTCCTGGCCCCCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 346 CCGCTTCAATGACCTCAACCGGAGAGCTGTGAACATGAGGCTTTCCGAGATGAGCTCCG 405
Db 301 CCGCTTCAATGACCTCAACCGGAGAGCTGTGAACATGAGGCTTTCCGAGATGAGCTCCG 360
QY 406 CAACCAATGCTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
Db 361 CAACCAATGCTGTGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 466 TGTTCGTGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
Db 421 TGTTCGTGAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 526 ACCTCTGAGGAGCTATAGAGGAGTGGGTTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 585
Db 481 ACCTCTGAGGAGCTATAGAGGAGTGGGTTGTTGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 586 AGCAGGATATTTGAGAGGAGATCTGTGTGAGGCTGTGAAGATGATCAAGAGAGAGAG 645
Db 541 AGCAGGATATTTGAGAGGAGATCTGTGTGAGGCTGTGAAGATGATCAAGAGAGAGAG 600
QY 646 CAGGTGTGTGCTCCCTTTGTGTGAAGCTGTGAAGAGAGAGAGAGAGAGAGAGAGAG 703
Db 601 CAGGTGTGTGCTCCCTTTGTGTGAAGCTGTGAAGAGAGAGAGAGAGAGAGAGAGAG 660
```



Qy	40	CTGGCAACCAAGCTGTGAGACGGATGACCTTCATCCTGCTCTTCTGCTCATGATG	460
Db	241	CTGGCAACCAAGCTGTGAGACGGATGACCTTCATCCTGCTCTTCTGCTCATGATG	300
Qy	461	CTTGTGTTGGTGGGCTCTCTCTGATGGACCTGTCTACCTGGAGTGGCAACCTGATGAG	520
Db	301	CTTGTGTTGGTGGGCTCTCTCTGATGGACCTGTCTACCTGGAGTGGCAACCTGATGAG	360
Qy	521	CGATG	525
Db	361	CGATG	365

```

RESULT 9
US-10-076-747-64
; Sequence 64, Application US/10076747
; Publication No. US20030180726A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karra, Kalpana
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and F
; FILE REFERENCE: DEX-0315
; CURRENT APPLICATION NUMBER: US/10/076,747
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,250
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/268,834
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-076-747-64

```

Query Match	30.5%;	Score 363.2;	DB 15;	Length 612;
Best Local Similarity	87.7%;	Pred. No. 1e-86;		
Matches 456;	Conservative	0;	Mismatches 53;	Indels 11; Gaps 5

Qy	666	CTTTGGTTAAACCTGGAGGGCAGAGGGAGACCTTTAGTCAG - CATTTCCAAAGGTGAGTG	720
Db	90	CTTTGGTTAAACCTGGAGGGAGAGGAGACCTTTAGTCAGACATTTCCAAAGGTGAGTG	149
Qy	721	GGTCCGTTGGTCCCGAGATACCTTTTAGTGGTATAGGGGCTTCGATTAAAGTGCAC - AAA	779
Db	150	GGTCCGTTGGTCCCGAGATACCTTTTAGTGGTATAGGGGCTTCGATTAAAGTGCACAAAA	209
Qy	780	ATCAGAGCAAGAAACGATGCGCTT - CCGATTTCTCATTCCTTTT - ATGCGCAGAG	836
Db	210	ATCAGAGCAAGAAACGATGCGCTTCCCGCATTTCTCATTCCTTTTATGCGCGAGAG	269
Qy	837	ATCTCAGCTGGATGCGAAACATGTTCCGATGCTGTGGAAGACATGCCGAGCTCTCTCTG	896
Db	270	ATCTCAGCTGGATGCGAAACATGTTCCGATGCTGTGGAAGACATGCCGAGCTCTCTCTG	329
Qy	897	CCTAGGAGCAAGACTTTGGGCTTAGGGCAGGTGCAAAAAATTCGACATTTTATAGACT	956
Db	330	CCTAGGAGCAAGACTTTGGGCTTAGGGCAGGTGCAAAAAATTCGACATTTTATAGACT	389
Qy	957	GTTTTTGGTTTATGATATATTTTATTTTATTTGCTACTTTATTTGTTAGACAGTGGTAGTG	1016
Db	390	GTTTTTGGTTTATGATATATTTTATTTTATTTGCTACTTTATTTGTTAGACAGTGGTAGTG	449
Qy	1017	GCATTC-----TATTTATTGGACCTTTTCAATCAATATGACTTTATAGTAAAAAATTTAAAAA	1076

Db 450 GCATTCCTCATTTATTATGGGGACCTTTCTCATATTAATAAGTATTATGCGCAAAAAAAAAA 509

Qy 1071 AAAAAAAAAAAAAAAAAAGCTGTGGGGGTTTACCCGGGGCCAAAGGGGGAGCCCGGGGGGAAT 1130

Db 510 ACAAATAAAAAAAAAAAGGCTGGGGGAAACCGGGGCCAAAGCCTGTTCGGGGGTGACAT 569

Qy 1131 GTGCTTCTCCGCCCAATTCGCCCAATTTTGTCAA 1170

Db 570 TGGTTTCCGGCCCAAAATTTCCAAATAATTTGGACAAA 609

RESULT 10  
US-10-027-632-174763/c  
; Sequence 174763, Application US/10027632  
; Publication No. US20020198371A1  
CENTRAL INFORMATION

```

? TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
? FILE OF INVENTION: Polymorphisms in the Human Genome
? FILE REFERENCE: 108827.129
? CURRENT APPLICATION NUMBER: US/10/027,632
? CURRENT FILING DATE: 2002-04-30
? PRIOR APPLICATION NUMBER: US 60/218, 006
? PRIOR FILING DATE: 2000-07-12
? PRIOR APPLICATION NUMBER: US 60/198, 676
? PRIOR FILING DATE: 2000-04-20
? PRIOR APPLICATION NUMBER: US 60/193, 493
? PRIOR FILING DATE: 2000-03-29
? PRIOR APPLICATION NUMBER: US 60/185, 218
? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/167, 363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156, 358
? PRIOR FILING DATE: 1999-09-28
? PRIOR APPLICATION NUMBER: US 60/146, 002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ. ID NOS: 325720
? SOFTWARE: fastseq for Windows Version 4.0
? SEQ. ID NO 174763
? LENGTH: 2940917
? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1) .. (2940917)
? OTHER INFORMATION: n = A,T,C or G
? US-10-027-632-174763

```

	Query Match	Best Local Similarity	6.48%;	Score 112.4;	DB 13;	Length 2940917;
			95.58%;	Pred. 10.4e-17;		
	Matches 161;	Conservative	0;	Mismatches 81;	Indels	Gaps 0;
QY	840	TCACGTGGATGCCCAACAATGTTCCGATGCGCTGGAGACATGCGGACGTCTCCCTGCTGCT				899
Db	2551564	TCTTCTCATAACCAACAATTTTCCGATACCTGTGTGCACATGTGGGGCTGTGACTTGCCT				2551505
QY	900	AGGAGCAGACGACTTGCGGCTTAGGCGCAGGTGGAAAAATTCGAGCTTTTATAGCACCTGTT				959
Db	2551504	AGGAGCCGAGAGTTAGTCGAGGGGAGGTCTAAAGAAATTTTAAACCTTTTATGATTTGTT				2551445
QY	960	TTTGTTTAAATGGAATATTTTAAATGAGCTACTTATATGTTTAGGACACAAGTGTAGTGGCA				1019
Db	2551444	TTTAACCGTATGGGAGGCTTTTGTGGCTACGATTAATCTTTGGGACTAGTCGTGTGACA				2551385
QY	1020	TTCTATTTATTTGACCTTTTCATATAATGATTTAGTATGAAAAAATAAAAAATAAAAA				1079
Db	2551384	TGTTAGATTTTATTTGCTTTTATTAAGTAAATTTTAAAGTAAGGATGCGCTGATTAAGAGA				2551325
QY	1080	AA 1081				
Db	2551324	AA 2551323				



```
RESULT 11
US-10-027-632-174763/c
; Sequence 174763, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174763
; LENGTH: 2940917
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2940917)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174763

Query Match      9.4%; Score 112.4; DB 16; Length 2940917;
Best Local Similarity 66.5%; Pred. No. 4.3e-17;
Matches 161; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY      840 TCAGCTGATGCCAATGTTCCGATGCTGTGAGAGATGCGGACGCTCTCCCTGCT 899
DB      2551564 TCTTTCATACCAACATTTTCGATACCTGTGCGATGCGGCGTCTGATGCGCT 2551505

QY      900 AGGAGACAGAGACTTGCGCTTACGCGAGGAGTGAAGAAAATTCAGACTTTTTCAGCACTGTT 959
DB      2551504 AGGTGACCGAGTATGCGAGGAGGAGGCTAAAGAAATTCAGACTTTTTCAGCACTGTT 2551445

QY      960 TTTGTTTAATGCTAATTTTATTTGCTTACTTTATTTAGGACAAAGTGTAGTGGCA 1019
DB      2551444 TTTAACCTGATGAGAGGTTTGTGCTTACGATTAATCTTTGGGACTTACGTCGAGCA 2551385

QY      1020 TTTGATTTATTTGACCTTTTCAATTAATGATTATTAAGTAAAGAAAAA 1079
DB      2551384 TGTTAAGTTTATTTGCTTTTATTAAGTAAATTTTAAGTAAAGATGCGTCAATAAGACA 2551325

QY      1080 AA 1081
DB      2551324 AA 2551323

RESULT 12
US-10-074-475-38
; Sequence 38, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Kaira, Kalpana
; APPLICANT: Cafferkey, Robert
```

```
APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-38

Query Match      6.3%; Score 75.6; DB 15; Length 461;
Best Local Similarity 62.9%; Pred. No. 1.2e-09;
Matches 117; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY      986 GCTACTTATTTAGTGAACAAGTGTAGTGCATTCATTTATTTGACCTTTTCATA 1045
DB      271 GCTTTTCCTTTTGTGTGCTCAAGTGTGCTGCATTTTAAATTTGTGAAATTAATAAC 330

QY      1046 AATGATTTAATGAAAAAAGCTGTGCGGCTTACCGCGG 1105
DB      331 TTTGTTATTTAGAAAAAAGCTGTGCGGCTTACCGCGG 390

QY      1106 CCAAAAGGCGGCGCGGCGGATGTGCTTCCGCCCAATTCGCCCAATTTTG 1165
DB      391 CCAAAAGCGGTGCGCGGCGGATGTGCTTCCGCCCAATTCGCCCAAAATTCG 450

QY      1166 AAAAA 1171
DB      451 AAAAA 456

RESULT 13
US-10-074-475-39
; Sequence 39, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Kaira, Kalpana
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; PRIOR FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 633
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-39

Query Match      6.3%; Score 75.6; DB 15; Length 633;
Best Local Similarity 62.9%; Pred. No. 1.5e-09;
Matches 117; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY      986 GCTACTTATTTAGTGAACAAGTGTAGTGCATTCATTTATTTGACCTTTTCATA 1045
DB      443 GCTTTTCCTTTTGTGTGCTCAAGTGTGCTGCATTTTAAATTTGTGAAATTAATAAC 502
```



Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1192	100.0	1192	6	AX431088	AX431088 Sequence
2	1033	86.7	1129	9	BC013587	BC013587 Homo sapi
3	792	66.4	1049	6	BD191093	BD191093 186 human
4	693	58.1	165110	9	AL450311	AL450311 Human DNA
5	1689	57.8	173341	2	AC021954	AC021954 Homo sapi
6	473	39.7	483	6	AR413227	AR413227 Sequence
7	473	31.7	483	6	BD108780	BD108780 EST and e
8	369	31.0	480	6	AX113082	AX113082 Sequence
9	311	26.1	450	11	G22561	G22561 human STR W
10	185	15.5	519	6	AR415427	AR415427 Sequence
11	185	15.5	519	6	BD110980	BD110980 EST and e
12	125	10.5	643	6	AX431087	AX431087 Sequence
13	38	3.2	68584	9	AC013065	AC013065 Homo sapi
14	38	3.2	118521	9	AC119049	AC119049 Homo sapi
15	38	3.2	118521	9	HPAC004514	HPAC004514 Homo sapi
16	38	3.2	118521	9	AC040159	AC040159 Homo sapi
17	38	3.2	153456	1	AC009311	AC009311 Homo sapi
18	37	3.1	193014	2	AC139357	AC139357 Homo sapi
19	36	3.0	201371	2	AC119162	AC119162 Mus muscu
20	36	3.0	243372	2	EX571714	EX571714 Dantio rer
21	36	3.0	328561	2	EX321879	EX321879 Dantio rer
22	35	2.9	148955	5	AY089465	AY089465 Drosophil
23	35	2.9	148955	3	EX004813	EX004813 Zebrafish
24	35	2.9	155287	2	EX284669	EX284669 Dantio rer
25	35	2.9	188180	2	AC023979	AC023979 Homo sapi
26	35	2.9	200041	2	EX571972	EX571972 Dantio rer
27	35	2.9	203948	10	AL627385	AL627385 Mouse DNA
28	35	2.9	233948	2	AC107741	AC107741 Mus muscu
29	34	2.9	436	9	BC062727	BC062727 Homo sapi
30	34	2.9	505	3	AF379640	AF379640 Heliothis
31	34	2.9	604	10	BC025044	BC025044 Mus muscu
32	34	2.9	1287	9	BC009221	BC009221 Homo sapi
33	34	2.9	1836	9	BC046243	BC046243 Homo sapi
34	34	2.9	2232	9	AK130049	AK130049 Homo sapi
35	34	2.9	2836	6	AR170129	AR170129 Sequence
36	34	2.9	2836	6	AR170130	AR170130 Sequence
37	34	2.9	2836	6	BD065859	BD065859 Novel cari
38	34	2.9	2836	6	BD065850	BD065850 Novel cari
39	34	2.9	3010	9	BC060865	BC060865 Homo sapi
40	34	2.9	43305	2	AC104240	AC104240 Homo sapi
41	34	2.9	18106	10	AF326737	AF326737 Mus muscu
42	34	2.9	109612	9	AL512503	AL512503 Human DNA
43	34	2.9	119456	2	AC084789	AC084789 Homo sapi
44	34	2.9	123619	10	AL663050	AL663050 Mouse DNA
45	34	2.9	123585	9	AC004707	AC004707 Homo sapi

proteins  
Patent: WO 0240535-A 31 23-MAY-2002;  
DIADEXUS INC (US)  
Location/Qualifiers  
1. 1192  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 100.0%; Score 1192; DB 6; Length 1192;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AAACCCCTGAGACCCCGGAGATGACGCGCCGAGAGCCCGGAGATGCGCCCTGACAGCA 60
Db 1 AAACCCCTGAGACCCCGGAGATGACGCGCCGAGAGCCCGGAGATGCGCCCTGACAGCA 60
QY 61 CAGCATCTACTCAGCGTGGTCACTCTGTGAACATCACTGACCTGCAAGCCTCCCTCAAT 120
Db 61 CAGCATCTACTCAGCGTGGTCACTCTGTGAACATCACTGACCTGCAAGCCTCCCTCAAT 120
QY 121 TTCTGTGACAGCCCATGAGGAGCCACAGCGCTGAGAGATGATGCGATCTTGGCCAA 180
Db 121 TTCTGTGACAGCCCATGAGGAGCCACAGCGCTGAGAGATGATGCGATCTTGGCCAA 180
QY 181 TGGGGAATGTGTGAGAGACGACGCGCGAGTGAAGACCACTACCAAGCACAAGAGG 240
Db 181 TGGGGAATGTGTGAGAGACGACGCGCGAGTGAAGACCACTACCAAGCACAAGAGG 240
QY 241 TAGCATCTCTGACAGAGCTTCTTCAATAGGGGCAATGATGCTCCCGCAGGGGGCTCTG 300
Db 241 TAGCATCTCTGACAGAGCTTCTTCAATAGGGGCAATGATGCTCCCGCAGGGGGCTCTG 300
QY 301 CCCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 CCCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 CAACCGGACAGTGTGAAATAGGAGCTTCCGAGTGGCATCTCGGCAACCATCTGTGA 420
Db 361 CAACCGGACAGTGTGAAATAGGAGCTTCCGAGTGGCATCTCGGCAACCATCTGTGA 420
QY 421 GCGGGAATCTCTGACAGAGCTTCTTCAATAGGGGCAATGATGCTCCCGCAGGGGGCTCTG 480
Db 421 GCGGGAATCTCTGACAGAGCTTCTTCAATAGGGGCAATGATGCTCCCGCAGGGGGCTCTG 480
QY 481 CCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 CCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 TAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 TAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 AGGGGATCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 AGGGGATCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 CCTTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 CCTTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 GGTTCGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 GGTTCGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 TCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 TCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 CAGGTGATGCAACATGATTCAGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 CAGGTGATGCAACATGATTCAGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
```

```
QY 901 GGGAGACAGACTTGGGCTTAGGGCAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 GGGAGACAGACTTGGGCTTAGGGCAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 TTGTTTATGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1020
Db 961 TTGTTTATGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1020
QY 1021 TCTATTATATGATCTTTTCAATTAATAGATTTAAGTAAAAAATATATATATATATATATATAT 1080
Db 1021 TCTATTATATGATCTTTTCAATTAATAGATTTAAGTAAAAAATATATATATATATATATATAT 1080
QY 1081 AAAAGAGCTGTGGGGTTACCGGGGGCCAAAGGGGGCCCGGGGGGAGATGATGCTCC 1140
Db 1081 AAAAGAGCTGTGGGGTTACCGGGGGCCAAAGGGGGCCCGGGGGGAGATGATGCTCC 1140
QY 1141 CCCCCCAATTCCTCCCAATTTTGAACAAAATGAAGAGACACAGGAACCC 1192
Db 1141 CCCCCCAATTCCTCCCAATTTTGAACAAAATGAAGAGACACAGGAACCC 1192
```

## RESULT 2

BC013587  
LOCUS  
DEFINITION  
Homo sapiens chromosome 10 open reading frame 35, mRNA (cDNA clone  
MGC:9596 IMAGE:389656), complete cds.

ACCESSION  
BC013587  
VERSION  
BC013587.1 GI:15488919  
KEYWORDS  
MGC.

SOURCE  
ORGANISM  
Homo sapiens (human)

REFERENCE  
1 (bases 1 to 1129)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Dere, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Schmeun, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Umed, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.U., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,  
Scherer, A., Schein, J.E., Jones, S.J. and Marz, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED  
12477932

REFERENCE  
2 (bases 1 to 1129)  
Strausberg, R.

AUTHORS  
Direct Submission  
Submitted (04-SEP-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
USA

REMARK  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [gcgaps-remail.nih.gov](mailto:gcgaps-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LINH)  
DNA Sequencing Arr: Sequencing Group at the Stanford Human Genome

Contact: (Dickson, Mark) mcd@paxil.stanford.edu  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R.M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>  
Series: IRAX Plate: 14 Row: 1 Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21897152.

# FEATURES

## Source

1. 1129

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:9596 IMAGE:389656"

/tissue\_type="Pancreas, epitheloid carcinoma"

/clone\_id="NIH\_MGC\_70"

/lab\_host="DH10B"

/note="Vector: PCMV-Sports"

## gene

1. 1129

/gene="C10orf35"

/db\_xref="LocusID:219738"

## CDS

160..525

/codon\_start=1

/product="chromosome 10 open reading frame 35"

/protein\_id="AAH13587.1"

/db\_xref="GI:15488920"

/db\_xref="LocusID:219738"

/translation="MYRIALANGIYDDDDPRVATTTOPRGSIFRGSFNRGAPRG  
GPRPQOAGARLGAOSPFNDLNRLQVNMSPFOMLGNHAYEPVTSILLFLMLMG  
VRELVLGVLYVSHLSOR"

# ORIGIN

Query Match 86.7%; Score 1033; DB 9; Length 1129;  
Best Local Similarity 99.9%; Pred. No. 0;

Matches 1089; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 2 AACCGCTGAGAGCCGCGGAGTGAACGCGCGGAGCCGCGGAGTGGCGCTTGCAGACAC 61
DB 1 AACCGCTGAGAGCCGCGGAGTGAACGCGCGGAGCCGCGGAGTGGCGCTTGCAGACAC 60
QY 62 AGCATCTACGAGGCTGCTCACTCTGTGAACATCTGACGACAGCTCCCTCAATT 121
DB 61 AGCATCTACGAGGCTGCTCACTCTGTGAACATCTGACGACAGCTCCCTCAATT 120
QY 122 TCTGTGACAGCCATCAAGGAGCCAGAGCGCTGGAGAGTGTGGGATCTTGGCCAAAT 181
DB 121 TCTGTGACAGCCATCAAGGAGCCAGAGCGCTGGAGAGTGTGGGATCTTGGCCAAAT 180
QY 182 GGGGAATCTGTGACAGAGCCAGAGCGCTGGAGAGTGTGGGATCTTGGCCAAAT 241
DB 181 GGGGAATCTGTGACAGAGCCAGAGCGCTGGAGAGTGTGGGATCTTGGCCAAAT 240
QY 242 AGCATCTCTGACAGAGCTTCTCAATAGAGGAGGAGTGTGGGATCTTGGCCAAAT 301
DB 241 AGCATCTCTGACAGAGCTTCTCAATAGAGGAGGAGTGTGGGATCTTGGCCAAAT 300
QY 302 CCCCAGAGAGAGAGAGTGTGGGATCTTGGGATCTTGGGATCTTGGGATCTTGGGATCTT 361
DB 301 CCCCAGAGAGAGAGAGTGTGGGATCTTGGGATCTTGGGATCTTGGGATCTTGGGATCTT 360
QY 362 AACCGAGAGCTGTGAACATGAGGCTTCCGAGTGGATCTCCGAGACATGCTGTGGAG 421
DB 361 AACCGAGAGCTGTGAACATGAGGCTTCCGAGTGGATCTCCGAGACATGCTGTGGAG 420
QY 422 CCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481
DB 421 CCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
QY 482 CTGGTGGCTTGTCTACTCTGTGTCCACCTGAGTACGCGGTGACCTTGAAGGCTGAT 541
DB 481 CTGGTGGCTTGTCTACTCTGTGTCCACCTGAGTACGCGGTGACCTTGAAGGCTGAT 540
```

```
QY 542 AGGAGTGGATTTGTGAGAGGAGTCTTGGGCTTGTGTGAGAGAGCATATTTGGA 601
DB 541 AGGAGTGGATTTGTGAGAGGAGTCTTGGGCTTGTGTGAGAGCATATTTGGA 600
QY 602 GGGATCTGTGTGTGCTTGAAGGTATGATCAAGAGGGAGACACAGTGTGTGTTTCC 661
DB 601 GGGATCTGTGTGTGCTTGAAGGTATGATCAAGAGGGAGACACAGTGTGTGTTTCC 660
QY 662 CTTTGTGTAAGCGTGAAGAGAGGAGACCTTGTGATTCAGATTTCCAAAGTGTGTG 721
DB 661 CTTTGTGTAAGCGTGAAGAGGAGACCTTGTGATTCAGATTTCCAAAGTGTGTG 720
QY 722 GTCCGTGTGTCCGAGATTAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 781
DB 721 GTCCGTGTGTCCGAGATTAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 782 CAGAGCAAGAAAGAGATGCTTCCCAATCTCTCAATCTTGTGTGTGTGTGTGTGTGT 841
DB 781 CAGAGCAAGAAAGAGATGCTTCCCAATCTCTCAATCTTGTGTGTGTGTGTGTGTGT 840
QY 842 AGCTGATGCAAGATGTTTCCAGATGCTGTGGAAGACATGCGAGTCTCTGTGCTAG 901
DB 841 AGCTGATGCAAGATGTTTCCAGATGCTGTGGAAGACATGCGAGTCTCTGTGCTAG 900
QY 902 GAGAGAGATCTTGGCTTGAAGGAGAGTGAAGAAATTCAGACTTTTGAAGCTGTTT 961
DB 901 GAGAGAGATCTTGGCTTGAAGGAGAGTGAAGAAATTCAGACTTTTGAAGCTGTTT 960
QY 962 TGTGTAAGTGAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1021
DB 961 TGTGTAAGTGAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 1022 CTTTGTGTGTGACCTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1081
DB 1021 CTTTGTGTGTGACCTTTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
QY 1082 AAAA 1085
DB 1081 AAAA 1084
```

RESULT 3  
BD191093 1049 bp DNA linear PAT 17-JUL-2003

LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

1 (bases 1 to 1049)  
Ruben, S.M., Rosen, C.A., Fischer, C.L., Soppet, D.R., Carter, K.C.,  
Bednarik, J.D., P., Endress, G.A., Yu, G.L., N.J., Feng, P., Young, P.E.,  
Greene, J.M., Ferris, A.M., Duan, R., Hu, J.S., Florence, K.A.,  
Olsen, H.S., Eder, R., Brewer, L.A., Moore, P.A., Shi, Y.,  
Lafleur, D.W., Li, Y., Zeng, Z., and Kyaw, H.

186 human secreted proteins  
Patent: JP 2002510192-A 57 02-APR-2002;  
HUMAN GENOME SCIENCES INC  
PN JP 2002510192-A/57

PD 02-APR-2002  
PF 06-MAR-1998 JP 1998538883  
PR 07-MAR-1997 US 60/040162, 07-MAR-1997 US 60/040333 PR  
07-MAR-1997 US 60/038621, 07-MAR-1997 US 60/040161 PR  
07-MAR-1997 US 60/040628, 07-MAR-1997 US 60/040334 PR  
07-MAR-1997 US 60/040336, 07-MAR-1997 US 60/040163 PR  
11-APR-1997 US 60/043580, 11-APR-1997 US 60/043568 PI  
KENNETH C CARTER, DANIEL P BEDNARIK, DANIEL R SOPPET, PI  
A ENDRESS, GUO LIANG  
PI YU, JIAN NI,  
PI PING FENG, PAUL E YOUNG, JOHN M GREENE, ANN

M. FERRE, ROXANNE DUAN,  
PI JING SHAN HU, KIMBERLY A FLORENCE, HENRIK  
S OLSEN, REINHARD EBERER,  
PI LAURIE A BREWER, PAUL A MOORE, YANGUO SHI, DAVID W LAFLEUR PI  
, YI LI, ZHI ZHEN ZENG,  
PI HLA KYAM  
PC C12N15/12, C12N5/10, C12N1/21, C07K14/47, C07K16/18, C12Q1/68, PC  
G01N33/50,  
CC Strandness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
Location/Qualifiers  
1. 1049  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

Query Match 66.4%; Score 792; DB 6; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

ORIGIN

Query Match 66.4%; Score 792; DB 6; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

46 TCGGCGCTGCAAGACAGCATCTACTAGGCTGGGTCACTCTGTGAACATCACTGACTG 105  
1 TCGGCGCTGCAAGACAGCATCTACTAGGCTGGGTCACTCTGTGAACATCACTGACTG 60  
106 CAAGCTCTCCCTCAATTTCTGGTGAAGCCCATCAAGGACCCACAGCGCTGGAGATGT 165  
61 CAAGCTCTCCCTCAATTTCTGGTGAAGCCCATCAAGGACCCACAGCGCTGGAGATGT 120  
166 GCGGATCTTGGCCAAATGAGGAAATGTCAGAGACAGACAGACCCCGAGTGAAGCACTA 225  
121 GCGGATCTTGGCCAAATGAGGAAATGTCAGAGACAGACAGACCCCGAGTGAAGCACTA 180  
226 CCAGCCCAAGAGAGTGAAGTCTCTGAGAGAGTCTTCAATAGGGGCCATGATCTCC 285  
181 CCAGCCCAAGAGAGTGAAGTCTCTGAGAGAGTCTTCAATAGGGGCCATGATCTCC 240  
286 CCAGAGGGGATCTGAGCCCGCCGAGAGAGAGAGTCCAGGCTGGTCTGCTCAGTC 345  
241 CCAGAGGGGATCTGAGCCCGCCGAGAGAGAGAGTCCAGGCTGGTCTGCTCAGTC 300  
346 CCCCTTCAATGACCTCAACCGGAGCTGAGTGAACATAGGGCTTCCGAGTGGATCTCG 405  
301 CCCCTTCAATGACCTCAACCGGAGCTGAGTGAACATAGGGCTTCCGAGTGGATCTCG 360  
406 CAACCATGCTGAGAGCGGAGCTGACCTCATCTGCTCTCTCTCTCATGATGCTGAG 465  
361 CAACCATGCTGAGAGCGGAGCTGACCTCATCTGCTCTCTCTCTCTCATGATGCTGAG 420  
466 TGTTCGTGGCTCTCTCTCTGTTGAGCTGCTGATCTGATCTCCACCTGATGAGCGGTG 525  
421 TGTTCGTGGCTCTCTCTCTGTTGAGCTGCTGATCTGATCTCCACCTGATGAGCGGTG 480  
526 ACCCTAGAGGCTATAGGGGATGAGTGTGTTGAGAGGAACTTCTGAGGCTTGTGAGT 585  
481 ACCCTAGAGGCTATAGGGGATGAGTGTGTTGAGAGGAACTTCTGAGGCTTGTGAGT 540  
586 AGCAGGATATTTGAGAGGAGATCTGCTGCTCTTGAAGGTATGATCAAGAGGAGCA 645  
541 AGCAGGATATTTGAGAGGAGATCTGCTGCTCTTGAAGGTATGATCAAGAGGAGCA 600  
646 CAGGTGTGTTCCTCTCTGTTGTTAAGGCTGAGGAGAGGAGAGGATAGT-CCAGCAT 704  
601 CAGGTGTGTTCCTCTCTGTTGTTAAGGCTGAGGAGAGGAGAGGATAGT-CCAGCAT 660  
705 TT-CCAAAGTGTGGGTGGTCCGTTGGTTCCGAGATACCTTTAGTGGTATGGGCTCTG 763  
661 TTCCCAAGTGTGGGTGGTCCGTTGGTTCCGAGATACCTTTAGTGGTATGGGCTCTG 720  
764 CATTAATGTGCTCAAAATGCTCAAAATGCTCAAAATGCTCAAAATGCTCAAAATGCTCA 823

Db 721 CATTAAGTGGCAGAAATCAGAGCAAGAGAGATGCCCTTCCCAATCTCTCATCTT 780  
QY 824 TTATGCCAGAAAGATCTCAGCTGATGCCAATGTTCCGATGCTGTGGAAGCATGCC 883  
Db 781 TTATGCCAGAAAGATCTCAGCTGATGCCAATGTTCCGATGCTGTGGAAGCATGCC 840  
QY 884 GACGTCTCTCTGCTTGAAGAGAGAGCTTGGGCTTGAAGAGAGAGAGAGAGAGAG 943  
Db 841 GACGTCTCTCTGCTTGAAGAGAGAGCTTGGGCTTGAAGAGAGAGAGAGAGAGAG 900  
QY 944 CTTTATGACACTGTTTGTGTTTATGAGTATTTTATTTATTTATTTATTTATTTAG 1003  
Db 901 CTTTATGACACTGTTTGTGTTTATGAGTATTTTATTTATTTATTTATTTATTTAG 960  
QY 1004 ACAAGTGTAGTGCATCTATTTATTTATTTATTTATTTATTTATTTATTTATTTAG 1063  
Db 961 ACAAGTGTAGTGCATCTATTTATTTATTTATTTATTTATTTATTTATTTATTTAG 1020  
QY 1064 AAAAAAAAAAAAAA 1077  
Db 1021 AAAAAAAAAAAAAA 1034

RESULT 4  
AL450311  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AL450311  
Human DNA sequence from clone RP11-343J3 on chromosome 10, complete  
sequence.  
AL450311  
GI:14626972  
HTG.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 165110)  
Howden, P.  
Direct Submission  
Submitted (12-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
On Jul 8, 2001 this sequence version replaced gi:14575291.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em, EMBL; SW,  
SWISSPROT; Tr, TrEMBL; Wp, WormBase; Information on the WormBase  
database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormbase This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 10, constructed by the Sanger Centre Chromosome 10  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr10  
RP11-343J3 is from the library RPCT-11.2 constructed by the group  
of Pieper de Jong. For further details see  
http://www.choi.org/bacpac/home.htm  
VECTOR: pBAC3.6

FEATURES  
This sequence is the entire insert of clone RP11-343J3. The true  
left end of clone RP11-24G20 is at 13995 in this sequence. The  
true right end of clone RP11-404C6 is at 6588 in this sequence.

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-343J3"
/clone_1b="RPC1-11.2"
repeat_region
7..147
/note="TH1C repeat: matches 2..142 of consensus"
151..293
/note="MIR repeat: matches 2..148 of consensus"
1136..1187
/note="26 copies 2 mer gt 98% conserved"
2960..3372
/note="AluSg repeat: matches 1..310 of consensus"
4144..4270
/note="AluSx repeat: matches 1..134 of consensus"
4319..4494
/note="AluSx repeat: matches 118..293 of consensus"
5662..6438
/note="L1MC repeat: matches 272..1095 of consensus"
6916..8080
/note="L1MC repeat: matches 1168..2367 of consensus"
9934..10087
/note="MIR repeat: matches 25..184 of consensus"
10353..10481
/note="MIR repeat: matches 65..194 of consensus"
11077..11372
/note="AluSx repeat: matches 1..294 of consensus"
11600..11639
/note="20 copies 2 mer tc 95% conserved"
12479..12608
/note="65 copies 2 mer at 90% conserved"
13433..13876
/note="MIR repeat: matches 1..466 of consensus"
13372..14576
/note="AluSg repeat: matches 1..301 of consensus"
14577..14633
/note="AluSg repeat: matches 243..299 of consensus"
14843..14943
/note="MER86 repeat: matches 9..109 of consensus"
15404..15613
/note="MIR repeat: matches 3..213 of consensus"
16626..16705
/note="MIR repeat: matches 60..140 of consensus"
16928..16993
/note="33 copies 2 mer gg 66% conserved"
17476..17559
/note="MER81 repeat: matches 2..114 of consensus"
17719..18069
/note="L1MC4 repeat: matches 7617..7977 of consensus"
18117..18345
/note="MIR repeat: matches 7..262 of consensus"
19169..19400
/note="MIR repeat: matches 5..239 of consensus"
19436..19627
/note="L1MC5 repeat: matches 7728..7917 of consensus"
19642..19716
/note="MIR repeat: matches 48..131 of consensus"
19969..20010
/note="21 copies 2 mer tg 100% conserved"
20623..20704
/note="41 copies 2 mer gt 85% conserved"
20738..20773
/note="U2 repeat: matches 1..36 of consensus"
20802..20863
/note="MIR repeat: matches 1..62 of consensus"
20889..21263
/note="MIR repeat: matches 188..541 of consensus"
21463..21618
/note="MIR repeat: matches 46..192 of consensus"
22019..22326
/note="AluB repeat: matches 1..306 of consensus"
22381..22564
repeat_region

repeat_region
/note="L2 repeat: matches 2453..2629 of consensus"
22896..23174
/note="MIR repeat: matches 117..413 of consensus"
23215..23346
/note="MIR repeat: matches 48..188 of consensus"
23388..23532
/note="L2 repeat: matches 2097..2230 of consensus"
23533..23857
/note="AluB repeat: matches 1..303 of consensus"
23838..24137
/note="L2 repeat: matches 1754..2097 of consensus"
24291..24581
/note="AluSx repeat: matches 1..300 of consensus"
24653..24850
/note="MIR repeat: matches 1..200 of consensus"
26189..26685
/note="CpG island"
29453..30918
/note="CpG island"
evidence-not_experimental
31084..31156
/note="MIR repeat: matches 65..138 of consensus"
31618..31876
/note="AluB repeat: matches 29..275 of consensus"
32767..32830
/note="MIR repeat: matches 76..139 of consensus"
33050..33178
/note="43 copies 3 mer tcc 72% conserved"
35112..36201
/note="CpG island"
evidence-not_experimental
36551..36604
/note="27 copies 2 mer ac 94% conserved"
37585..38254
/note="L1P15 repeat: matches 5480..6157 of consensus"
38265..38561
/note="AluSx repeat: matches 1..300 of consensus"
39214..39423
/note="MIR repeat: matches 3..219 of consensus"
39489..39801
/note="AluSx repeat: matches 1..292 of consensus"
40050..40189
/note="MIR repeat: matches 1..144 of consensus"
41057..41290
/note="MIR repeat: matches 8..240 of consensus"
41296..41460
/note="L2 repeat: matches 2569..2730 of consensus"
41506..41944
/note="L2 repeat: matches 1916..2416 of consensus"
42388..42698
/note="L2 repeat: matches 1448..1779 of consensus"
44193..44579
/note="TH1C repeat: matches 1..371 of consensus"
44600..44956
/note="L1P16 repeat: matches 90..445 of consensus"
45240..45500
/note="MER58A repeat: matches 37..97 of consensus"
45798..45909
/note="L1P1 repeat: matches 90..192 of consensus"
46826..46871
/note="23 copies 2 mer gt 97% conserved"
50684..50969
/note="AluSg repeat: matches 1..287 of consensus"
50980..51291
/note="AluSx repeat: matches 1..312 of consensus"
52222..52519
/note="AluSx repeat: matches 1..300 of consensus"
54065..54260
/note="L1M4 repeat: matches 3865..4055 of consensus"
54261..54432
/note="FAM repeat: matches 2..167 of consensus"
54433..54629
repeat_region
```





Consensus quality: 161190 bases at least Q40  
Consensus quality: 166837 bases at least Q30  
Consensus quality: 168995 bases at least Q20  
Insert size: 176000; agarose-fp  
Quality coverage: 3.7 in Q20 bases; agarose-fp  
Quality coverage: 3.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1574: contig of 1574 bp in length  
1575 1674: gap of 100 bp  
1675 3043: contig of 1369 bp in length  
3044 3143: gap of 100 bp  
3144 5889: contig of 2746 bp in length  
5890 5990: gap of 100 bp  
5990 8979: contig of 2990 bp in length  
8980 9079: gap of 100 bp  
9080 13674: contig of 4595 bp in length  
13675 13775: gap of 100 bp  
13775 18831: contig of 5057 bp in length  
18832 18931: gap of 100 bp  
18932 23526: contig of 4595 bp in length  
23527 23627: gap of 100 bp  
23627 27386: contig of 3760 bp in length  
27387 27486: gap of 100 bp  
27487 32572: contig of 5086 bp in length  
32573 32672: gap of 100 bp  
32673 38632: contig of 5960 bp in length  
38633 38732: gap of 100 bp  
38733 43735: contig of 5003 bp in length  
43736 43835: gap of 100 bp  
43836 49020: contig of 5185 bp in length  
49021 49120: gap of 100 bp  
49121 53660: contig of 4540 bp in length  
53661 53760: gap of 100 bp  
53761 59544: contig of 5784 bp in length  
59545 59644: gap of 100 bp  
59645 66395: contig of 6751 bp in length  
66396 66495: gap of 100 bp  
66496 74696: contig of 8201 bp in length  
74697 74796: gap of 100 bp  
74797 81326: contig of 6530 bp in length  
81327 81426: gap of 100 bp  
81427 89446: contig of 8020 bp in length  
89447 89546: gap of 100 bp  
89547 96459: contig of 6913 bp in length  
96460 96559: gap of 100 bp  
96560 106498: contig of 9939 bp in length  
106499 106598: gap of 100 bp  
106599 116986: contig of 10388 bp in length  
116987 117086: gap of 100 bp  
117087 128890: contig of 11804 bp in length  
128891 128990: gap of 100 bp  
128991 147290: contig of 18300 bp in length  
147291 147390: gap of 100 bp  
147391 173341: contig of 25951 bp in length.

FEATURES

source  
1. 173341  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/map="10"  
/clone="RP11-57E12"  
/clone\_id="RPC1-11 Human Male BAC"  
misc\_feature  
1. 1574

misc\_feature /note="assembly\_fragment"  
1675..3043  
/note="assembly\_fragment"  
3144..5889  
/note="assembly\_fragment"  
5990..8979  
/note="assembly\_fragment"  
9080..13674  
/note="assembly\_fragment"  
13775..18831  
/note="assembly\_fragment"  
18932..23526  
/note="assembly\_fragment"  
23627..27386  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
27487..32572  
/note="assembly\_fragment"  
32673..38632  
/note="assembly\_fragment"  
38733..43735  
/note="assembly\_fragment"  
43836..49020  
/note="assembly\_fragment"  
49121..53660  
/note="assembly\_fragment"  
53761..59544  
/note="assembly\_fragment"  
59645..66395  
/note="assembly\_fragment"  
66496..74696  
/note="assembly\_fragment"  
74797..81326  
/note="assembly\_fragment"  
81427..89446  
/note="assembly\_fragment"  
89547..96459  
/note="assembly\_fragment"  
96560..106498  
/note="assembly\_fragment"  
106599..116986  
/note="assembly\_fragment"  
117087..128890  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right  
128991..147290  
/note="assembly\_fragment"  
147391..173341  
/note="assembly\_fragment"

Query Match 57.8%; Score 689; DB 2; Length 173341;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 809; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 254 CAGAGCTTCTCAATAGAGGGCCATGTCCTCCCGAGGGGGTCTGGGCCCCGACGAG 313  
DB 161089 CAGAGCTTCTCAATAGAGGGCCATGTCCTCCCGAGGGGGTCTGGGCCCCGACGAG 161030  
QY 314 CAGGAGGTGCCAGGCTGGGTCTGCTCAGTCCCTTCATAGACTCAACCGGAGGCTG 373  
DB 161022 CAGGAGGTGCCAGGCTGGGTCTGCTCAGTCCCTTCATAGACTCAACCGGAGGCTG 160970  
QY 374 GTGACATGGGGCTTTCGCGAGTGCATCTCGGCAACCATGCTGTGAGACCGGTGACCTCC 433  
DB 160969 GTGACATGGGGCTTTCGCGAGTGCATCTCGGCAACCATGCTGTGAGACCGGTGACCTCC 160910  
QY 434 ATCTGCTCTCTTCTCTGCTCATGATGCTGTGTTCTGACCTCTCTGTTGACCTT 493  
DB 160909 ATCTGCTCTCTTCTCTGCTCATGATGCTGTGTTCTGACCTCTCTGTTGACCTT 160850  
QY 494 GTCTACTGTGTCTCCACCTGATGACGCGTGAACCTGTGAGGCTGATAGGGGTGGGTTT 553

Db 16049 GTCTACCTGCTGCTCCACCTGAGTCAAGGCTGACCTCTGAGGCTGATGAGGGTGGGTTT 160790  
QY 554 GTTGAAGAGGAGCTTGTCTGAGCTTGTGTGAGAGCAGGCAATATTGGAGGGATCTGTG 613  
Db 160789 GTTGAAGAGGAGCTTGTCTGAGCTTGTGTGAGAGCAGGCAATATTGGAGGGATCTGTG 160730  
QY 614 GTGCTTGAAGAGTATGATCAGAGAGGGAGCAGAGTGTGTGTTCCCTTTGTGTAA 673  
Db 160729 GTGCTTGAAGAGTATGATCAGAGAGGGAGCAGAGTGTGTGTTCCCTTTGTGTAA 160670  
QY 674 CGTAGAGCAGAGGAGAGCTTGTGAGAGCTTCCAAAGTGTGGTGGTCCGTTGGTTC 733  
Db 160669 CGTAGAGCAGAGGAGAGCTTGTGAGAGCTTCCAAAGTGTGGTGGTCCGTTGGTTC 160610  
QY 734 CCGAGATACCTTTAGTGTGTATGGGGCTTGCAATTAAGTGCACAAATCAGACAGAGAA 793  
Db 160609 CCGAGA-ACCTTTAGTGTGTATGGGGCTTGCAATTAAGTGCACAAATCAGACAGAGAA 160551  
QY 794 GCGATGCGCTTCCCAATTCCTCTCAATCCTTTATGCGGAGAGATCTCAGCTGATGCCA 853  
Db 160550 GCGATGCGCTTCCCAATTCCTCTCAATCCTTTATGCGGAGAGATCTCAGCTGATGCCA 160491  
QY 854 ACATGTTCCGATGCTGTGAGAGACATGCCAGCTCTCTCTGCTTGAAGGAGAGAGACTT 913  
Db 160490 ACATGTTCCGATGCTGTGAGAGACATGCCAGCTCTCTCTGCTTGAAGGAGAGAGACTT 160431  
QY 914 GCGCTTGAAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 973  
Db 160430 GCGCTTGAAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 160371  
QY 974 ATATTTTATGCTGACTTTATTTGTTTGAAGAGAGTGTGAGAGAGAGAGAGAGAGAG 1033  
Db 160370 ATATTTTATGCTGACTTTATTTGTTTGAAGAGAGTGTGAGAGAGAGAGAGAGAG 160311  
QY 1034 ACCTTTTCAATTAATAGATTAAATTAATTAATTAATTAATTAATTAATTAATTA 1063  
Db 160310 ACCTTTTCAATTAATAGATTAAATTAATTAATTAATTAATTAATTAATTAATTA 160281

RESULT 6  
AR413227 483 bp DNA linear PAT 18-DEC-2003  
LOCUS AR413227  
DEFINITION Sequence 864 from patent US 6639063.  
ACCESSION AR413227  
VERSION AR413227.1 GI:40168337  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 483)  
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.  
TITLE EST's and encoded human proteins  
JOURNAL Patent: US 6639063-A 864 28-OCT-2003;  
FEATURES  
source 1..483  
Location/Qualifiers  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 39.7%; Score 473; DB 6; Length 483;  
Best Local Similarity 100.0%; Pred. No. 3,4e-278; Indels 0; Gaps 0;  
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 131 TTCTGTGACAGCCATTCAGGAGACCCAGCCCTGGAGAGATGTGGATCTTGGCCAA 190  
QY 181 TGGGGAATTCGTGACAGACAGACACCCCGAGTAGAGACCACTACCAACCAACAAGAG 240  
Db 191 TGGGGAATTCGTGACAGACAGACACCCCGAGTAGAGACCACTACCAACCAACAAGAG 250  
QY 241 TAGCATTTCTGACAGACCTTCTTCATAGGGGCAATGTGTCTCCCGAGGGGCTCTGG 300  
Db 251 TAGCATTTCTGACAGACCTTCTTCATAGGGGCAATGTGTCTCCCGAGGGGCTCTGG 310  
QY 301 CCCCCGACAGAGCAGAGAGTGCAGAGCTGGTGTGCTCTCAGTCCCTTCAATGACT 360  
Db 311 CCCCCGACAGAGCAGAGAGTGCAGAGCTGGTGTGCTCTCAGTCCCTTCAATGACT 370  
QY 361 CAACCGGAGCTGTGAAACATAGGGCTTTCCGAGTGGATCTCGGCAACCATGTGTGA 420  
Db 371 CAACCGGAGCTGTGAAACATAGGGCTTTCCGAGTGGATCTCGGCAACCATGTGTGA 430  
QY 421 GCCGATGACCTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 473  
Db 431 GCCGATGACCTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 483

RESULT 7  
BD108780 483 bp DNA linear PAT 18-SEP-2002  
LOCUS BD108780  
DEFINITION EST and encoded human protein.  
ACCESSION BD108780  
VERSION BD108780.1 GI:23203598  
KEYWORDS JP 2002010789-A/857.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 483)  
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-E.  
TITLE EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 857 15-JAN-2002;  
COMMENT OS Homo sapiens (human)  
PN JP 2002010789-A/857  
PD 15-JAN-2002 JP 2000280989  
PF 07-AUG-2000 JP 60/147499  
PI 05-AUG-1999 US 60/147499  
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21  
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC EST and encoded human protein  
FH Key Location/Qualifiers  
FT CDS 171..482.  
Location/Qualifiers  
1..483  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 39.7%; Score 473; DB 6; Length 483;  
Best Local Similarity 100.0%; Pred. No. 3,4e-278; Indels 0; Gaps 0;  
Matches 473; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 121 TTGTGTGACGCCCATATGAGGACCCACAGGCGCTGGAGAGATGTGCGATCTTGGCCAA 180
Db 131 TTGTGTGACGCCCATATGAGGACCCACAGGCGCTGGAGAGATGTGCGATCTTGGCCAA 190
QY 181 TGGGGAATGTGTGAGAGACCAACCCCGATGAGGACCACTACCAAGCACCACAGAG 240
Db 191 TGGGGAATGTGTGAGAGACCAACCCCGATGAGGACCACTACCAAGCACCACAGAG 250
QY 241 TAGCATCTCTGACAGAGCTTCTTCATATAGGGGCGCATGTGCTCCCGAGGGGCTCTGG 300
Db 251 TAGCATCTCTGACAGAGCTTCTTCATATAGGGGCGCATGTGCTCCCGAGGGGCTCTGG 310
QY 301 CCCCCGACAGACAGGAGGAGGTCAGAGGCTGGTGTCTGCTCACTGCCCTTCAATAGCT 360
Db 311 CCCCCGACAGACAGGAGGAGGTCAGAGGCTGGTGTCTGCTCACTGCCCTTCAATAGCT 370
QY 361 CAACCCGACAGCTGTGTGAACATGGGCTTCCGAGTGGCATCTGGGCAACCATCTGTGA 420
Db 371 CAACCCGACAGCTGTGTGAACATGGGCTTCCGAGTGGCATCTGGGCAACCATCTGTGA 430
QY 421 GCGGCTGACCTCCATCTCTGCTCTTCTCTGCTCATGATGCTTGTGCTG 473
Db 431 GCGGCTGACCTCCATCTCTGCTCTTCTCTGCTCATGATGCTTGTGCTG 483

RESULT 8
AX113082 550 bp DNA linear PAT 01-MAY-2001
LOCUS AX113082
DEFINITION Sequence 5 from Patent WO0127276.
ACCESSION AX113082
VERSION AX113082.1 GI:13939514
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Harlocker, S.L., Dillon, D.C. and Xu, J.
Dna sequences from breast tumor and uses thereof
Patent: WO 0127276-A 5 19-APR-2001;
CORIXA CORPORATION (US)

FEATURES
source 1..550
location/Qualifiers
1..550
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
/misc_feature 1..550
/note="n = A,T,C or G"

ORIGIN
Query Match 31.0%; Score 369; DB 6; Length 550;
Best Local Similarity 100.0%; Pred. No. 2.4e-214;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 CCGCGCTGACAGACAGCATCTACTAGCGGTGCTCACTCTGTGAACATCACTGACTGC 106
Db 1 CCGCGCTGACAGACAGCATCTACTAGCGGTGCTCACTCTGTGAACATCACTGACTGC 60

QY 107 AAGCCTCCCTAATTTCTGTGTCAGCCCATGAGGACCCACAGGCTGGAGAGATGTG 166
Db 61 AAGCCTCCCTAATTTCTGTGTCAGCCCATGAGGACCCACAGGCTGGAGAGATGTG 120

QY 167 CGGATCTTGGCAATGGGGAATGTGTGAGACAGACACCCCGAGTGAAGACCACTACC 226
Db 121 CGGATCTTGGCAATGGGGAATGTGTGAGACAGACACCCCGAGTGAAGACCACTACC 180

QY 227 CAGCCACCAAGAGTAGATCTCTGACAGAGCTTCTTCATATAGGGGCGCATGTGCTCC 286
Db 181 CAGCCACCAAGAGTAGATCTCTGACAGAGCTTCTTCATATAGGGGCGCATGTGCTCC 240

QY 287 CCAAGGGGCTCTGCGCCCGGACAGAGAGGAGGTCGAGGCTGGTGTGCTCACTGC 346
Db 241 CCAAGGGGCTCTGCGCCCGGACAGAGAGGAGGTCGAGGCTGGTGTGCTCACTGC 300

```

```

QY 347 CCTTCATAGACTCAACCGGAGCTGTGAACATGGGCTTCCGAGTGGCATCTCGGC 406
Db 301 CCTTCATAGACTCAACCGGAGCTGTGAACATGGGCTTCCGAGTGGCATCTCGGC 360
QY 407 AACCATGCT 415
Db 361 AACCATGCT 369

RESULT 9
G22561/c 480 bp DNA linear STS 31-MAY-1996
LOCUS G22561
DEFINITION human STS WI-12709, sequence tagged site.
ACCESSION G22561
VERSION G22561.1 GI:1342887
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 480)
Hudson, T.
Whitehead Institute/MIT Center for Genome Research, Physically
Mapped STS
Unpublished (1995)

JOURNAL
COMMENT
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu

Primer A: AAAGGTCACATTAATGAAATGCC
Primer B: GCTTAGGGGAGGTGAAAAA
STS size: 125
PCR Profile:
Presack:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 mM
Tag Polymerase: 0.025 units/ul
Total Vol: 20 ul
Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 9.3

Derived from dbEST (genbank accession H49529).
Location/Qualifiers
1..480
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="431.2 CR from top of Chr10 linkage group"
25..149
primer_bind 25..149
primer_bind complement(130..149)

ORIGIN
Query Match 26.1%; Score 311; DB 11; Length 480;
Best Local Similarity 100.0%; Pred. No. 9.4e-179;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 733 CCCGAGATACCTTTTAAAGTGTATGAGGCGCTGCAATTAAGTGCACAAATGCAAGCAAGAA 792  
DB 332 CCCGAGATACCTTTTAAAGTGTATGAGGCGCTGCAATTAAGTGCACAAATGCAAGCAAGAA 273  
QY 793 AGCGATGCCCTTCCCAATCTCTCAATCTCTTTTATGCGAGAGATCTCGAGTATGCC 852  
DB 272 AGCGATGCCCTTCCCAATCTCTCAATCTCTTTTATGCGAGAGATCTCGAGTATGCC 213  
QY 853 AACATGTTCCGATGCTGTGAGAGACATGCGAGCTCTCTGCTGCTAGGAGAGAGACT 912  
DB 212 AACATGTTCCGATGCTGTGAGAGACATGCGAGCTCTCTGCTGCTAGGAGAGAGACT 153  
QY 913 TGGGCTTAGGGGAGAGTGAAGAAATTCAGACTTTTATGACACTGTTTGTATTAATGG 972  
DB 152 TGGGCTTAGGGGAGAGTGAAGAAATTCAGACTTTTATGACACTGTTTGTATTAATGG 93  
QY 973 TATATTTTATGCTACTTATGTTTATGTTAGACAGAGTGTAGTGGCACTTCTATTTATGT 1032  
DB 92 TATATTTTATGCTACTTATGTTTATGTTAGACAGAGTGTAGTGGCACTTCTATTTATGT 33  
QY 1033 GACCTTTTCAA 1043  
DB 32 GACCTTTTCAA 22

## RESULT 10

LOCUS AR415427 519 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 3064 from patent US 6639063.  
ACCESSION AR415427  
VERSION AR415427.1 GI:40170537  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 519)  
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.  
TITLE EST's and encoded human proteins  
JOURNAL Patent: US 6639063-A 3064 28-OCT-2003;  
FEATURES  
Source Location/Qualifiers  
1..519  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 15.5%; Score 185; DB 6; Length 519;  
Best Local Similarity 99.6%; Pred. No. 2.1e-101; Indels 0; Gaps 0;  
Matches 235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 ACAAGAGATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACAGCTCCCTC 117  
DB 116 ACAAGAGATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACAGCTCCCTC 175  
QY 118 AATTTCTGTGACGCCCATCAGAGGACCCACAGCGCTCGAGAGATGTTGGATCTTGGC 177  
DB 176 AATTTCTGTGACGCCCATCAGAGGACCCACAGCGCTCGAGAGATGTTGGATCTTGGC 235  
QY 178 CAATGGGGAATCGTGAGAGACGACGACCCCGAGTGAAGACCACTAACCCAGCCACAAG 237  
DB 236 CAATGGGGAATCGTGAGAGACGACGACCCCGAGTGAAGACCACTAACCCAGCCACAAG 295  
QY 238 AGGTAGCATCTCTGACAGAGCTTCTTCAATTAGGGGCGCATGTGCTTCCCGAGGGG 293  
DB 296 AGGTAGCATCTCTGACAGAGCTTCTTCAATTAGGGGCGCATGTGCTTCCCGAGGGG 351

RESULT 11  
LOCUS BD110980 519 bp DNA linear PAT 18-SEP-2002  
DEFINITION EST and encoded human protein.  
ACCESSION BD110980  
VERSION BD110980.1 GI:23205798

KEYWORDS JP 2002010789-A/3057.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 519)  
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.  
TITLE EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 3057 15-JAN-2002;  
COMMENT GENSET CORP  
OS Homo sapiens (human)  
PN JP 2002010789-A/3057  
PD 15-JAN-2002  
PF 07-AUG-2000 JP 2000280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC EST and encoded human protein  
CC Key Location/Qualifiers  
FH 93..431.  
FT CDS Location/Qualifiers  
1..519  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 15.5%; Score 185; DB 6; Length 519;  
Best Local Similarity 99.6%; Pred. No. 2.1e-101; Indels 0; Gaps 0;  
Matches 235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 ACAAGAGATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACAGCTCCCTC 117  
DB 116 ACAAGAGATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACAGCTCCCTC 175  
QY 118 AATTTCTGTGACGCCCATCAGAGGACCCACAGCGCTCGAGAGATGTTGGATCTTGGC 177  
DB 176 AATTTCTGTGACGCCCATCAGAGGACCCACAGCGCTCGAGAGATGTTGGATCTTGGC 235  
QY 178 CAATGGGGAATCGTGAGAGACGACGACCCCGAGTGAAGACCACTAACCCAGCCACAAG 237  
DB 236 CAATGGGGAATCGTGAGAGACGACGACCCCGAGTGAAGACCACTAACCCAGCCACAAG 295  
QY 238 AGGTAGCATCTCTGACAGAGCTTCTTCAATTAGGGGCGCATGTGCTTCCCGAGGGG 293  
DB 296 AGGTAGCATCTCTGACAGAGCTTCTTCAATTAGGGGCGCATGTGCTTCCCGAGGGG 351

## RESULT 12

AX431087 643 bp DNA linear PAT 28-JUN-2002  
LOCUS AX431087  
DEFINITION Sequence 30 from Patent WO0240335.  
ACCESSION AX431087  
VERSION AX431087.1 GI:21656064  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Recipon,H., Liu,C., Salceda,S., Sun,Y., Caffery,R. and Macina,R.A.  
TITLE Compositions and methods relating to ovary specific genes and proteins

JOURNAL Patent: WO 0240335-A 30 23-MAY-2002;  
DIADEXUS INC (US)  
FEATURES  
Source Location/Qualifiers  
1..643  
/organism="Homo sapiens"

```

/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match      10.5%; Score 125; DB 6; Length 643;
Best Local Similarity 100.0%; Pred. No. 1.5e-64;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 828 GCCGAGAAATCTCAGCTGATGCCAATGTTCCGATGCTGTGGAAGACATGCCGAG 887
Db 275 GCCGAGAAATCTCAGCTGATGCCAATGTTCCGATGCTGTGGAAGACATGCCGAG 334
Cy 888 TCTCTCTGCTTGGAGAGAGAGACTTGGGCTTGGGAGAGTGAATAATTCAGACTTT 947
Db 335 TCTCTCTGCTTGGAGAGAGAGACTTGGGCTTGGGAGAGTGAATAATTCAGACTTT 944
Cy 948 TTATG 952
Db 395 TTATG 399

RESULT 13
LOCUS AC019069 68584 bp DNA linear PRI 21-FEB-2002
DEFINITION Homo sapiens BAC clone RP11-259N19 from 2, complete sequence.
ACCESSION AC019069
VERSION AC019069.8 GI:18042390
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 68584)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 68584)
AUTHORS Harris, A. and Haglund, K.
TITLE The sequence of Homo sapiens BAC clone RP11-259N19
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 68584)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 68584)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 68584)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 3, 2002 this sequence version replaced gi:16259101.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0259N19
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

```

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

**MAPPING INFORMATION:**  
Mapping information for this clone was provided by Dr. John D.  
McPherson, Department of Genetics, Washington University, St. Louis  
MO. For additional information about the map position of this  
sequence, see <http://genome.wustl.edu/gsc>

**SOURCE INFORMATION:**  
The RPCT-11 human BAC library was made from the blood of one male  
donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Firengen, B.,  
Tateno, M., Catanese, J. V. and de Jong, P. J. (1998) An improved  
approach for construction of bacterial artificial chromosome  
libraries. *Genomics* 51:1-8. The clone may be obtained either from  
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong  
and coworkers at <http://www.chori.org>  
VECTOR: pBACe3.6

**NEIGHBORING SEQUENCE INFORMATION:**  
The clone sequenced to the left is RP11-372J12, 2000 bp overlap.  
The clone sequenced to the right is RP11-588I4. Actual end is at  
base position 68584 of RP11-259N19.

Polymorphisms have been identified between AC007387 and AC019069.

```

FEATURES
    source
        1..68584
            /organism="Homo sapiens"
            /mol_type="Genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="2"
            /map="2"
            /clone="RP11-259N19"
            /clone_id="RPCT-11"
            /size="153..300"
            /rpt_family="L2"
            /rpt_size="280..332"
            /rpt_family="CA)n"
            /rpt_size="1455..1675"
            /rpt_family="L1"
            /rpt_size="1755..2031"
            /rpt_family="ERVL"
            /rpt_size="2032..2323"
            /rpt_family="Alu"
            /rpt_size="2324..2556"
            /rpt_family="ERVL"
            /rpt_size="4046..4551"
            /rpt_family="L2"
            /rpt_size="4555..4859"
            /rpt_family="MER2_type"
            /rpt_size="4873..5205"
            /rpt_family="MER2_type"
            /rpt_size="5212..5251"
            /rpt_family="(CA)n"
            /rpt_size="5255..5833"
            /rpt_family="L2"
            /rpt_size="6960..7028"
            /rpt_family="L1"
            /rpt_size="7029..7327"
            /rpt_family="Alu"
            /rpt_size="7328..7393"
            /rpt_family="L1"
            /rpt_size="7433..7625"
            /rpt_family="L1"
            /rpt_size="7626..7931"
            /rpt_family="Alu"
            /rpt_size="7631..7650"
            /rpt_size="7631..7650"

```

[illegible]

AUTHORS

Loftus, B.J., Kim, U.J., Sneddon, V.P., Kalush, F., Brandon, R., Fuhrmann, J., Mason, T., Crosby, M.J., Barnstead, M., Cronin, L., Deslattes, Mays, A., Cao, Y., Xu, R.X., Kang, H.L., Mitchell, S., Eichler, B.E., Harris, P.C., Venter, J.C. and Adams, M.D.  
Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q  
Genomics 60 (3), 285-308 (1999)

JOURNAL

99425270

REFERENCE

2 (bases 1 to 118521)

AUTHORS

Adams, M.D., Loftus, B.J., Zhou, L., Crosby, M., Fuhrmann, J., Mason, T.M., Brandon, R., Kim, U.J., Kerlavage, A.R. and Venter, J.C.

JOURNAL

Unpublished

REFERENCE

3 (bases 1 to 118521)

AUTHORS

Adams, M.D. and Loftus, B.J.

JOURNAL

Submitted (31-MAR-1998)

REFERENCE

4 (bases 1 to 118521)

AUTHORS

Adams, M.D. and Loftus, B.J.

JOURNAL

Submitted (25-JUL-1998)

REFERENCE

Medical Center Dr., Rockville, MD 20850, USA

COMMENT

On Jul 25, 1998 this sequence version replaced gi:3212841. Address all correspondence to: Mark Adams The Institute for Genomic Research 9712 Medical Center Dr. Rockville, MD 20850, USA e-mail address: humgen@tigr.org. The orientation of the sequence is from 5p6 end to 17 end. Genes were identified by a combination of five methods including: XGRAIL (available by anonymous ftp from arthur.epm.ornl.gov), GeneFinder (Phil Green, University of Washington), Genscan (Chris Burge <http://genomic.stanford.edu/~chris/GENSCANW.html>) searches of the complete sequence against a peptide database, and the Human gene Index database at TIGR (<http://www.tigr.org/cdb/hgi/hgi.html>). Genes without peptide homology having spliced EST hits are termed 'Unknown gene product'. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Location/Qualifiers

FEATURES

source

1. 118521  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/map="16q21/22"  
/clone="A-954B10"  
53412..53582  
/note="42758, D11S1053, Chr. -, Homo sapiens"  
/db\_xref="dbSTS:G28812"  
53720..53897  
/note="7821, STS1-CSRL-31b6-uA/CSRL-31b6-uZ, Chr. -, Homo sapiens"  
/db\_xref="dbSTS:G02335"  
53536..56067  
/note="7617, STS1-CSRL-30a3-uA/CSRL-30a3-uZ, Chr. -, Homo sapiens"  
/db\_xref="dbSTS:G02131"

ORIGIN

Query Match 3.2%; Score 38; DB 9; Length 118521;  
Best Local Similarity 100.0%; Pred.No. 8.2e-11;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1048 TAGATTAAAGTAAAAAAAAAAAAAAAAAAAAA 1085  
Dp 27928 TAGATTAAAGTAAAAAAAAAAAAAAAAAAAAA 27965

Search completed: August 22, 2004, 06:19:10  
Job time : 4693 secs

**This Page Blank (uspto)**



OM nucleic - nucleic search, using sw model

Run on: August 22, 2004, 03:29:28 ; Search time 514 Seconds

(without alignments)  
9851.852 Million cell updates/sec

Title: US-10-001-885-31

Perfect score: 1192  
Sequence: 1 aaacgcgcgcgaagcgcgcg.....gaagaggaacacaggaacc 1192

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 212409041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

- 1: geneseqn1808.\*
- 2: geneseqn1908.\*
- 3: geneseqn2000.\*
- 4: geneseqn2001.\*
- 5: geneseqn2001b.\*
- 6: geneseqn2002.\*
- 7: geneseqn2003a.\*
- 8: geneseqn2003b.\*
- 9: geneseqn2003c.\*
- 10: geneseqn2004.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1192	100.0	1192	6	ABN87820 Human ova
2	903	75.8	1108	6	ABL90057 Human ova
3	792	66.4	1049	2	AAV59558 Human sec
4	792	66.4	1049	6	AB873545 Human cDN
5	792	66.4	1049	8	ACD82668 CDNA sequ
6	562	47.1	562	6	ABZ11212 Human pol
7	369	31.0	550	4	AA82505 Human bre
8	326	27.3	365	7	ABX92251 Human ova
9	235	19.7	517	7	ABX74673 Human cDN
10	195	16.4	612	7	ABX92250 Human ova
11	161	13.5	1165	5	AA864820 DNA encod
12	159	13.3	388	5	AA864819 DNA encod
13	157	13.2	587	5	AA864821 DNA encod
14	155	10.5	643	6	ABN87819 Human ova
15	38	3.2	473	4	AA116223 Probe #61
16	38	3.2	473	4	AA116223 Probe #61
17	38	3.2	473	4	AA116223 Probe #61
18	38	3.2	473	4	ABAS8927 Human foe
19	38	3.2	473	4	AA188644 Probe #73
20	38	3.2	473	4	AA137326 Probe #60
21	38	3.2	473	4	ABAS7125 Probe #55
22	38	3.2	473	4	ABAS7125 Probe #55
23	38	3.2	473	4	AAK32829 Human bon

24	38	3.2	473	4	AAK31436 Human bon
25	38	3.2	473	4	AAK07089 Human bra
26	38	3.2	473	4	AAK05810 Human bra
27	38	3.2	473	4	AB832554 Human liv
28	38	3.2	473	4	AB831118 Human liv
29	38	3.2	473	6	AB806190 Human gen
30	38	3.2	473	6	AB807633 Human gen
31	34	2.9	2836	2	AAV40745 Human gen
32	34	2.9	2836	2	AAV40744 C. felis
33	34	2.9	2836	2	AAV40744 C. felis
34	34	2.9	3288	4	AAAD05613 Human sec
35	34	2.9	3393	4	AAAD05582 Human sec
36	33	2.8	653	6	ABO57023 Human col
37	33	2.8	1021	6	ABK86007 Rainbow t
38	33	2.8	1256	6	AA84605 Human ZFS
39	33	2.8	43064	7	ACC80906 Human CYP
40	33	2.8	96595	9	ADA02936 Human BRA
41	33	2.8	96595	9	ADB72674 Human BRA
42	33	2.8	96595	9	ADC85416 Human Bra
43	32	2.7	157	4	AA860109 Human can
44	32	2.7	157	4	AB186599 Human ova
45	32	2.7	473	6	ABV95508 Human pan

ALIGNMENTS

RESULT 1  
ABN87820  
ID ABN87820 standard; cDNA; 1192 BP.  
XX  
AC ABN87820;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE Human ovary specific nucleic acid SEQ ID NO:31.  
XX  
KW Human; ovary specific nucleic acid; OSNA; ovary specific protein; OSP;  
KW ovary specific gene; OSO; ovarian cancer; immune response; metastasis;  
KW chromosome 10; gene; se.  
XX  
XX Homo sapiens.  
OS  
XX WO200240535-A2.  
PN  
XX 23-MAY-2002.  
PD  
XX 20-NOV-2001; 2001WO-US045011.  
PF  
XX 20-NOV-2000; 2000US-0252061P.  
PR  
XX 27-NOV-2000; 2000US-0253257P.  
XX  
PA (DIAD-) DIADEXUS INC.  
XX  
PI Salceda S, Macina RA, Recipon H, Cafierkey R, Sun Y, Liu C;  
XX WPI, 2002-471617/50.  
DR  
XX New ovary specific genes and proteins, useful as a vaccine for treating  
XX patients with ovarian cancer, or for diagnosing and monitoring the  
XX presence and metastases of ovarian cancer in a patient.  
PT  
PT Claim 1; Page 173-174; 260pp; English.  
PS  
PS ABN87790 to ABN87882 represent human ovary specific nucleic acid (OSNA)  
PS sequences, and AB879297 to AB879370 represent ovary specific protein  
PS (OSP) sequences from the present invention. OSNA and OSP sequences have  
PS cytostatic activity, and can be used in vaccine production and gene  
PS therapy. An antibody that specifically binds to an OSP can be used for  
PS treating a patient with ovarian cancer, particularly for inducing an  
PS immune response against the ovarian cancer cell expressing the OSNA or  
PS OSP. The OSNA and OSPs can also be used for diagnosing and monitoring  
PS the presence and metastases of ovarian cancer in a patient

XX Sequence 1192 BP; 262 A; 298 C; 349 G; 283 T; 0 U; 0 Other;

Query Match 100.0%; Score 1192; DB 6; Length 1192;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AAACCGCTGAGAGCCGCGGAGTGAAGCGCCGCGAGAGCCGAGAGTGGCGCTTCAGACA 60
Db 1 AAACCGCTGAGAGCCGCGGAGTGAAGCGCCGCGAGAGCCGAGAGTGGCGCTTCAGACA 60

QY 61 CAGCATCTACTCAGCGGTGGGTCACTTGTGAACATCACTGACTGCAAGCTTCTTCAT 120
Db 61 CAGCATCTACTCAGCGGTGGGTCACTTGTGAACATCACTGACTGCAAGCTTCTTCAT 120

QY 121 TTCTGTGTCAGGCCCATCAGGAGACCCACAGCGCTGGAGAGATGATGCGGATCTTGSCAA 180
Db 121 TTCTGTGTCAGGCCCATCAGGAGACCCACAGCGCTGGAGAGATGATGCGGATCTTGSCAA 180

QY 181 TGGGGAATCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 TGGGGAATCGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240

QY 241 TAGATTCCTCGACAGGCTTCTTCAATAGGGGCAATGATGCTCCCGCAGGGGGTCTCTG 300
Db 241 TAGATTCCTCGACAGGCTTCTTCAATAGGGGCAATGATGCTCCCGCAGGGGGTCTCTG 300

QY 301 CCCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 CCCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 361 CAACCGGACAGTGTGAACATGAGGCTTCCGAGTGGCACTTGGGCAACAGTGTGTGA 420
Db 361 CAACCGGACAGTGTGAACATGAGGCTTCCGAGTGGCACTTGGGCAACAGTGTGTGA 420

QY 421 GCGGTGACCTCCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480
Db 421 GCGGTGACCTCCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480

QY 481 CCTGTGTGCTGTGTCTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db 481 CCTGTGTGCTGTGTCTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540

QY 541 TAGGGGTGGGTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 TAGGGGTGGGTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

QY 601 AGGGGATCTGATGTGCTTGAAGTATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 AGGGGATCTGATGTGCTTGAAGTATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

QY 661 CCTTTGTGTTAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 CCTTTGTGTTAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720

QY 721 GATCGGTGTGTTCCCGAGATCTTTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 GATCGGTGTGTTCCCGAGATCTTTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780

QY 781 TCGAGAGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 TCGAGAGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840

QY 841 CAGGTGATGCAATGTTCCAGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 CAGGTGATGCAATGTTCCAGATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

QY 901 GGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 GGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960

QY 961 TTGTTTAATGATATTTTATTTATGCTACTTTATTTTATGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 TTGTTTAATGATATTTTATTTATGCTACTTTATTTTATGAGAGAGAGAGAGAGAGAGAG 1020
```

```

Db 961 TTGTTTAATGATATTTTATTTATGCTACTTTATTTTATGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 TCTATTTATGTCAGCTTTTCAATTAATAGATTTTAAATTAATTAATTAATTAATTAATTA 1080
Db 1021 TCTATTTATGTCAGCTTTTCAATTAATAGATTTTAAATTAATTAATTAATTAATTAATTA 1080
QY 1081 AAAAAGCTGTGGGGGTTTACCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 AAAAAGCTGTGGGGGTTTACCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 GCGCCCAATTCGCCCAATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1192
Db 1141 GCGCCCAATTCGCCCAATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1192

RESULT 2
AB:90057
ID ABL90057 standard; cDNA; 1108 BP.
XX
AC ABL90057;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polymuclotide SEQ ID NO 619.
XX
KW Cytosolic; immunosuppressive; neotropic; neuroprotective; antiviral;
KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
KW vulnereary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200190304-R2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR MPI; 2002-122018/16.
XX
DR P-PSDB; ABB89648.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 4; SEQ ID NO 619; 2081bp + Sequence Listing; English.
XX
XX
The invention relates to novel genes (ABL9449-ABL90853) and proteins
(CC (ABR89040-ABR90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
```

[illegible]

PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047616P.  
PR 23-MAY-1997; 97US-0047617P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047632P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 06-JUN-1997; 97US-0048974P.  
PR 13-JUN-1997; 97US-0049610P.  
PR 08-JUL-1997; 97US-0051928P.  
PR 16-JUL-1997; 97US-0052874P.  
PR 18-AUG-1997; 97US-0055724P.  
PR 22-AUG-1997; 97US-0056530P.  
PR 22-AUG-1997; 97US-0056631P.  
PR 22-AUG-1997; 97US-0056632P.  
PR 22-AUG-1997; 97US-0056637P.  
PR 22-AUG-1997; 97US-0056638P.  
PR 22-AUG-1997; 97US-0056642P.  
PR 22-AUG-1997; 97US-0056643P.  
PR 22-AUG-1997; 97US-0056644P.  
PR 22-AUG-1997; 97US-0056845P.  
PR 22-AUG-1997; 97US-0056846P.  
PR 22-AUG-1997; 97US-0056872P.  
PR 22-AUG-1997; 97US-0056874P.  
PR 22-AUG-1997; 97US-0056875P.  
PR 22-AUG-1997; 97US-0056876P.  
PR 22-AUG-1997; 97US-0056877P.  
PR 22-AUG-1997; 97US-0056878P.  
PR 22-AUG-1997; 97US-0056879P.  
PR 22-AUG-1997; 97US-0056880P.  
PR 22-AUG-1997; 97US-0056881P.  
PR 22-AUG-1997; 97US-0056882P.  
PR 22-AUG-1997; 97US-0056884P.  
PR 22-AUG-1997; 97US-0056886P.  
PR 22-AUG-1997; 97US-0056887P.  
PR 22-AUG-1997; 97US-0056888P.  
PR 22-AUG-1997; 97US-0056889P.  
PR 22-AUG-1997; 97US-0056892P.  
PR 22-AUG-1997; 97US-0056893P.  
PR 22-AUG-1997; 97US-0056894P.  
PR 22-AUG-1997; 97US-0056903P.  
PR 22-AUG-1997; 97US-0056908P.  
PR 22-AUG-1997; 97US-0056909P.  
PR 22-AUG-1997; 97US-0056910P.  
PR 22-AUG-1997; 97US-0056911P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057669P.  
PR 05-SEP-1997; 97US-0057761P.  
PR 12-SEP-1997; 97US-0058785P.  
PR 02-OCT-1997; 97US-0061060P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Ruben SM, Rosen CA, Fischer CL, Soppet DR, Carter KC,  
PI Bednarik DP, Endress GA, Yu G, Ni J, Feng P, Young PG, Greene JM,  
PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA,  
PI Moore PA, Shi Y, Lathier DW, Li Y, Zeng Z, Kyaw H,  
XX  
XX WPI: 1998-506364/43.  
XX P-PSDB; AAW74778.  
XX  
XX New isolated human genes and the secreted polypeptide(s) they encode -  
PT

PT disorders, immune diseases, inflammation or blood disorders.  
XX  
XX Claim 1, Page 280, 721p; English.  
XX  
XX This sequence represents a nucleic acid molecule designated Gene 48 from  
CC the human CDNA clone HRC4174 (deposited as clone ATCC 97899 and ATCC  
CC 209045) which encodes a secreted human protein. The gene can be used to  
CC generate fusion proteins by linking to the gene to a human immunoglobulin  
CC portion (e.g. AAV59502) for increasing the stability of the fused  
CC protein as compared to the human protein only. The invention relates to  
CC 186 novel genes and their fragments (nucleic acid sequences: AAV59511-  
CC V59812; amino acid sequences AAW74731-W75026) which are useful for  
CC preventing, treating or ameliorating medical conditions e.g. by protein  
CC or gene therapy. Also, pathological conditions can be diagnosed by  
CC determining the amount of the new polypeptides in a sample or by  
CC determining the presence of mutations in the new polynucleotides.  
CC Specific uses are described for each of the 186 polynucleotides, based on  
CC which tissues they are most highly expressed in (see AAV59511 for  
XX described uses).  
SQ Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;  
Query Match 66.4%; Score 792; DB 2; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 5.1e-302;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
QY 46 TCGCCCTGCGACACACACATCTACTAGCGTGGGTACCTCTGTGAACATCACTGACTG 105  
DB 1 TCGCCCTGCGACACACACATCTACTAGCGTGGGTACCTCTGTGAACATCACTGACTG 60  
QY 106 CAAGCCTCCCTCAATTTGTGTGACGCCATCAGAGACCCACAGCGCTGGAGATGTGT 165  
DB 61 CAAGCCTCCCTCAATTTGTGTGACGCCATCAGAGACCCACAGCGCTGGAGATGTGT 120  
QY 166 GCGGATCTTGGCCAAATGGGGAAATGTGACAGACGACGCCCGAGTGAAGACCACTAC 225  
DB 121 GCGGATCTTGGCCAAATGGGGAAATGTGACAGACGACGCCCGAGTGAAGACCACTAC 180  
QY 226 CGAGCACCAAGAGGTGACATCTCTCCACAGAGCTTCTTCAATAGGGCCATGCTGTCC 285  
DB 181 CGAGCACCAAGAGGTGACATCTCTCCACAGAGCTTCTTCAATAGGGCCATGCTGTCC 240  
QY 286 CCGAGGGGGTCTTGGCCCGCCCGCAGACAGACGACGATCCAGGCTGGGTCTGCTAGTC 345  
DB 241 CCGAGGGGGTCTTGGCCCGCCCGCAGACAGACGATCCAGGCTGGGTCTGCTAGTC 300  
QY 346 CCCCTTCAATGACCTCAACCGGCGAGCTGTGACATGATGGCTTCCGAGTGGCATCTCGG 405  
DB 301 CCCCTTCAATGACCTCAACCGGCGAGCTGTGACATGATGGCTTCCGAGTGGCATCTCGG 360  
QY 406 CAACCATGCTGTGAGCGCGGTGACCTTCATCTGCTCTCTTCCGCTCAATGATCTTGG 465  
DB 361 CAACCATGCTGTGAGCGCGGTGACCTTCATCTGCTCTCTTCCGCTCAATGATCTTGG 420  
QY 466 TGTTCGTGGCTCTCTCTGTGTGGCTGTGTACCTGTGTGTCCCACTGATCAGCGGTG 525  
DB 421 TGTTCGTGGCTCTCTCTGTGTGGCTGTGTACCTGTGTGTCCCACTGATCAGCGGTG 480  
QY 526 ACCTCTAGAGGCTGATAGGGGTGGGTTTGTGAAGAGGACCTTGCGGGCTTGCTGTGAG 585  
DB 481 ACCTCTAGAGGCTGATAGGGGTGGGTTTGTGAAGAGGACCTTGCGGGCTTGCTGTGAG 540  
QY 586 AGCAGGCAATATTGGAGGGGATCTGTGTGCTTGAAGGTATGATCAGAGGGGACCA 645  
DB 541 AGCAGGCAATATTGGAGGGGATCTGTGTGCTTGAAGGTATGATCAGAGGGGACCA 600  
QY 646 CAGGTGTGTGTTCCCTTGTGTGTAAGGTGAGGACGAGAGAGAGTGTAGT-CCAGCAT 704  
DB 601 CAGGTGTGTGTTCCCTTGTGTGTAAGGTGAGGACGAGAGAGAGTGTAGTCCAGCAT 660  
QY 705 TT-CCAAAGTGTGGGTGGTCCGTGTGTTCCGAGATCTTTAGGTGTATGGGGCTTG 763

PR	11-APR-1997	97US-0043671P
PR	11-APR-1997	97US-0043672P
PR	11-APR-1997	97US-0043674P
PR	23-MAY-1997	97US-0047492P
PR	23-MAY-1997	97US-0047500P
PR	23-MAY-1997	97US-0047501P
PR	23-MAY-1997	97US-0047502P
PR	23-MAY-1997	97US-0047503P
PR	23-MAY-1997	97US-0047561P
PR	23-MAY-1997	97US-0047562P
PR	23-MAY-1997	97US-0047563P
PR	23-MAY-1997	97US-0047585P
PR	23-MAY-1997	97US-0047586P
PR	23-MAY-1997	97US-0047587P
PR	23-MAY-1997	97US-0047588P
PR	23-MAY-1997	97US-0047589P
PR	23-MAY-1997	97US-0047590P
PR	23-MAY-1997	97US-0047591P
PR	23-MAY-1997	97US-0047592P
PR	23-MAY-1997	97US-0047593P
PR	23-MAY-1997	97US-0047594P
PR	23-MAY-1997	97US-0047595P
PR	23-MAY-1997	97US-0047596P
PR	23-MAY-1997	97US-0047597P
PR	23-MAY-1997	97US-0047615P
PR	23-MAY-1997	97US-0047616P
PR	23-MAY-1997	97US-0047618P
PR	23-MAY-1997	97US-0047622P
PR	23-MAY-1997	97US-0047633P
PR	06-JUN-1997	97US-0048964P
PR	06-JUN-1997	97US-0048974P
PR	13-JUN-1997	97US-0049610P
PR	08-JUL-1997	97US-0051926P
PR	16-AUG-1997	97US-0052874P
PR	22-AUG-1997	97US-0056430P
PR	22-AUG-1997	97US-0056631P
PR	22-AUG-1997	97US-0056632P
PR	22-AUG-1997	97US-0056635P
PR	22-AUG-1997	97US-0056637P
PR	22-AUG-1997	97US-0056642P
PR	22-AUG-1997	97US-0056644P
PR	22-AUG-1997	97US-0056845P
PR	22-AUG-1997	97US-0056846P
PR	22-AUG-1997	97US-0056847P
PR	22-AUG-1997	97US-0056848P
PR	22-AUG-1997	97US-0056849P
PR	22-AUG-1997	97US-0056850P
PR	22-AUG-1997	97US-0056851P
PR	22-AUG-1997	97US-0056852P
PR	22-AUG-1997	97US-0056853P
PR	22-AUG-1997	97US-0056854P
PR	22-AUG-1997	97US-0056855P
PR	22-AUG-1997	97US-0056856P
PR	22-AUG-1997	97US-0056857P
PR	22-AUG-1997	97US-0056858P
PR	22-AUG-1997	97US-0056859P
PR	22-AUG-1997	97US-0056903P
PR	22-AUG-1997	97US-0056908P
PR	22-AUG-1997	97US-0056909P

22-AUG-1997; 97US-0056910P.  
PR 22-AUG-1997; 97US-0056911P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057659P.  
PR 05-SEP-1997; 97US-0057761P.  
PR 12-SEP-1997; 97US-0058785P.  
PR 02-OCT-1997; 97US-0061060P.  
PR 06-MAR-1998; 98WO-US004493.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PI Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC,  
PI Bednarik DR, Endress GA, Yu G, Ni J, Feng P, Young PR, Greene JM,  
PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA;  
PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
XX MPI, 2002-634796/68.  
DR P-PSDB; ABG95227.  
XX  
XX New isolated human secreted protein for diagnosing, preventing, treating  
PT or ameliorating medical conditions and used as a food additive or  
PT preservative.  
PS  
PS Example 1; SEQ ID NO 58; 129pp; English.  
XX

The invention relates to an isolated protein that is one of 186 human  
CC secreted proteins, given in the specification, encoded by one of 309 cDNA  
CC sequences also given in the specification. The protein is used in a  
CC pharmaceutical composition used to prevent, treat or ameliorate a medical  
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. Disorders which are diagnosed or treated include  
CC autoimmune diseases e.g. Rheumatoid arthritis, hyperproliferative  
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC e.g. corneal infection. The polypeptides can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,  
CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
CC components. The present sequence represents a cDNA derived from a gene  
CC encoding one of the novel human secreted proteins of the invention. Note:  
CC This sequence did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docid=6420526b1  
XX

XX Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

Query Match 66.4%; Score 792; DB 6; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 5,1e-302;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 46 TCGGGCGCTGAGACAGACAGATCTACTCAGCGTGGGTCACTGTGTAACATCACTGACTG 105  
DB 1 TCGGGCGCTGAGACAGACAGATCTACTCAGCGTGGGTCACTGTGTAACATCACTGACTG 60  
QY 106 CAAGCTTCCTCAATTTCTGTCAGAGCCATCAAGGAGCCCAAGCGCTGGAGAGATGGT 165  
DB 61 CAAGCTTCCTCAATTTCTGTCAGAGCCATCAAGGAGCCCAAGCGCTGGAGAGATGGT 120  
QY 166 GCGGATTTGGCCCAATGGGGAATCGTGCAGAGCGAGACCCCGAGTGGAGCCACTGAC 225  
DB 121 GCGGATTTGGCCCAATGGGGAATCGTGCAGAGCGAGACCCCGAGTGGAGCCACTGAC 180  
QY 226 CCAAGCCCAAGAGAGTGAATTCCTCGACAGAGCTTTCTTCAATAGGGGCCATGATGCTCC 285  
DB 181 CCAAGCCCAAGAGAGTGAATTCCTCGACAGAGCTTTCTTCAATAGGGGCCATGATGCTCC 240  
QY 286 CCAAGGGGGTCTGGGGCCCCCGCAGCAGCAGAGAGTCCAGGGCTGGTGGTCTGATGTC 345

DB 241 CCAAGGGGGTCTGGGGCCCCCGCAGCAGCAGAGAGTCCAGGCTGGATGCTGCTGATC 300  
QY 346 CCCCTTCAATACCTCAACCGGGAGCTGTGTGAATAGGGCTTTCCGAGTGGCATCTCGG 405  
DB 301 CCCCTTCAATACCTCAACCGGGAGCTGTGTGAATAGGGCTTTCCGAGTGGCATCTCGG 360  
QY 406 CAACCATGCTGTGAGAGCGCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 465  
DB 361 CAACCATGCTGTGAGAGCGCGGTGACCTTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
QY 466 TGTTCGTGGCT 525  
DB 421 TGTTCGTGGCT 480  
QY 526 ACCTCTGAGGCTGATAGGGGTGGGTTGTGAGAGGAGACTTGTCTGGGCTTGTGTGAG 585  
DB 481 ACCTCTGAGGCTGATAGGGGTGGGTTGTGAGAGGAGACTTGTCTGGGCTTGTGTGAG 540  
QY 586 AGCAGGCAATTTGGAGGGGATCTGTGTGCTCTTGAAGTATATCAAGAGGGGACCA 645  
DB 541 AGCAGGCAATTTGGAGGGGATCTGTGTGCTCTTGAAGTATATCAAGAGGGGACCA 600  
QY 646 CAGGTGTGTGTTCCCTTTGTGTTAAGCGTGAAGGAGAGGAGCTTGTAGT-CCAGCAT 704  
DB 601 CAGGTGTGTGTTCCCTTTGTGTTAAGCGTGAAGGAGAGGAGCTTGTAGTCCAGCAT 660  
QY 705 TT-CCAAAGTGTGGTGGTCCGTGGTTCGAGATCTTTTNGTGTATGGGGCTTG 763  
DB 661 TTCCAAAGTGTGGTGGTCCGTGGTTCGAGATCTTTTNGTGTATGGGGCTTG 720  
QY 764 CATTAAGTGGCAAAATCAAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 823  
DB 721 CATTAAGTGGCAAAATCAAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780  
QY 824 TTATCCGAGAGAGATCTCAGCTGATGCAATGTTCCGATGCTGTGAGAGCATGCC 883  
DB 781 TTATCCGAGAGAGATCTCAGCTGATGCAATGTTCCGATGCTGTGAGAGCATGCC 840  
QY 884 GACGCTCTCTCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 943  
DB 841 GACGCTCTCTCTGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
QY 944 CTTTCTTTCAGCTGTTTCTTCTTAAATGATATTTTATTTATTTATTTATTTATTTATTTAG 1003  
DB 901 CTTTCTTTCAGCTGTTTCTTCTTAAATGATATTTTATTTATTTATTTATTTATTTATTTAG 960  
QY 1004 ACAAGTGTAGTGGCATTTATTTATTTGAGCTTTTCAATTAATTAATTTAATTAATAA 1063  
DB 961 ACAAGTGTAGTGGCATTTATTTATTTGAGCTTTTCAATTAATTAATTTAATTAATAA 1020  
QY 1064 AAAAAAAAAAAAAA 1077  
DB 1021 AAAAAAAAAAAAAA 1034

RESULT 5  
ACD82688  
ID ACD82688 standard; cDNA, 1049 BP.

XX ACD82688;  
XX  
XX 22-SEP-2003 (first entry)

CDNA sequence #48 containing coding region of a human secreted protein.

Human, secreted protein, hyperproliferative disorder; leukaemia;  
Breast cancer; wound; reproductive disorder; blood-related disorder;  
Haemophilia; thrombocytopenia; immunodeficiency; thymic hypoplasia;  
Wiskott-Aldrich syndrome; autoimmune disorder; multiple sclerosis;  
Kw graft-versus-host disease; Hashimoto's thyroiditis; allergy; asthma;  
viral infection; bacterial infection; fungal infection; AIDS; sepsis;  
renal disorder; kidney failure; cardiovascular disorder; cytostatic;  
antineoplastic; cerebral ischaemia; congenital heart defect.



PI Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Ferris AM,  
PI Dian DR, Hu J, Florence KA, Olsen HS, Fischer CU, Ebner R,  
PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
XX WPI: 2003-521800/49.  
DR P-PSDB; ABO34421.

XX New genes and its encoded prostate cancer antigen proteins, useful for  
PT preventing, treating, ameliorating or diagnosing e.g. prostate cancers,  
PT thymic hypoplasia, multiple sclerosis, AIDS, angina pectoris or cerebral  
PT ischemia.

PS Claim 4; SEQ ID NO 58; 260bp; English.

XX The present invention relates to the isolation of novel human secreted  
CC proteins and the polynucleotide sequences encoding them. The invention  
CC also discloses vectors, host cells, antibodies, and recombinant methods  
CC for producing human secreted proteins. The polypeptide and polynucleotide  
CC sequences for the secreted proteins are useful for preventing, treating,  
CC ameliorating or diagnosing medical conditions such as hyperproliferative  
CC disorders (e.g. leukemia or breast cancers), wounds, reproductive  
CC disorders, blood-related disorders (e.g. haemophilia or  
CC thrombocytopenia), immunodeficiencies (e.g. Wiskott-Aldrich syndrome or  
CC thymic hypoplasia), autoimmune disorders (e.g. graft-versus-host disease,  
CC multiple sclerosis or Hashimoto's thyroiditis), allergies (e.g. asthma),  
CC viral or bacterial or fungal infections (e.g. AIDS or sepsis), renal  
CC disorders (e.g. kidney failure), cardiovascular disorders (e.g. angina  
CC pectoris, cerebral ischaemia or congenital heart defects), respiratory  
CC disorders, neurological disorders (e.g. Alzheimer's disease or  
CC Parkinson's disease), and inflammations (e.g. Crohn's disease). The  
CC polynucleotide or polypeptide may also be used as vaccine adjuvants.  
CC ACB82841:ACB82950 encode human secreted proteins or their fragments.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site at [seqdata.uspto.gov/psipdsidentity.html](http://seqdata.uspto.gov/psipdsidentity.html)

XX Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;

XX Query Match 66.4%; Score 792; DB 8; Length 1049;

XX Best Local Similarity 99.8%; Pred. No. 5,1e-302;

XX Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 46 TCGGCGCTGACAGACAGCATCTACTGAGTGGTCACTCTGTGAACATCACTGACTG 105  
DB 1 TCGGCGCTGACAGACAGCATCTACTGAGTGGTCACTCTGTGAACATCACTGACTG 60  
QY 106 CAAAGCTCCCTCAATTTCTGGTGAAGCCCATCAGGAGCCACAGCGCTGGAGAGATGGT 165  
DB 61 CAAAGCTCCCTCAATTTCTGGTGAAGCCCATCAGGAGCCACAGCGCTGGAGAGATGGT 120  
QY 166 GCGGATCTTGGCCCATGAGGAAATCTGTCAGAGACAGAGACCCCGAGTGAAGCACTAC 225  
DB 121 GCGGATCTTGGCCCATGAGGAAATCTGTCAGAGACAGAGACCCCGAGTGAAGCACTAC 180  
QY 226 CCAGCCACCAAGAGAGTGAATCTCTGACAGAGCTTCTCATATAGGGGCCCATGATGCTCC 285  
DB 181 CCAGCCACCAAGAGAGTGAATCTCTGACAGAGCTTCTCATATAGGGGCCCATGATGCTCC 240  
QY 286 CCGAGGGGGTCTGGCCCCCGCCAGACAGACAGAGTCCAGAGTGGGTGCTGCTCAATC 345  
DB 241 CCGAGGGGGTCTGGCCCCCGCCAGACAGACAGAGTCCAGAGTGGGTGCTGCTCAATC 300  
QY 346 CCGCTTCAATGAGCTCAACCGGAGCTGTGAACATGAGCTTCCGAGTGGGATCTCGG 405  
DB 301 CCGCTTCAATGAGCTCAACCGGAGCTGTGAACATGAGCTTCCGAGTGGGATCTCGG 360  
QY 406 CAAACATGCTGTGAGCCGGTGAATCTCTGCTCTCTTCTGCTCAATGATGCTTGG 465  
DB 361 CAAACATGCTGTGAGCCGGTGAATCTCTGCTCTCTTCTGCTCAATGATGCTTGG 420  
QY 466 TGTTCGTGAGCTCTCTCTGCTGTGAGCTTGTCTATAGCTGAGTCCCACTGATGAGGGG 525  
DB 421 TGTTCGTGAGCTCTCTCTGCTGTGAGCTTGTCTATAGCTGATGAGGGG 480

QY 526 ACCTGTGAGGGCTGATATGAGGGTGGTGGTTGTTGAGAGGAGACTGTGAGGCTTGGTGTAG 585  
DB 481 ACCTGTGAGGGCTGATATGAGGGTGGTGGTTGTTGAGAGGAGACTGTGAGGCTTGGTGTAG 540  
QY 586 AGCAGGCAATATTTGAGGGGATCTGTTGGTGGCTTTGAAGTATGATCAAGAGGGACCA 645  
DB 541 AGCAGGCAATATTTGAGGGGATCTGTTGGTGGCTTTGAAGTATGATCAAGAGGGACCA 600  
QY 646 CAGGTGTGTGTTCCCTTGTGTGAAGCGGAGGAGGAGGAGAGCTTGTAGT-CCAGCAT 704  
DB 601 CAGGTGTGTGTTCCCTTGTGTGAAGCGGAGGAGGAGGAGAGCTTGTAGTCCAGCAT 660  
QY 705 TT-CCAAAGTGTGGGTGGTCCGTTGGTTCGAGATACCTTTTGTGTGTGGGGCTTG 763  
DB 661 TTCCAAAGTGTGGGTGGTCCGTTGGTTCGAGATACCTTTTGTGTGTGGGGCTTG 720  
QY 764 CATTAAGTGACAAATTCAGAGCAAGAAAGCGATGCCCTTCCCAATCTCTCAATCCTT 823  
DB 721 CATTAAGTGACAAATTCAGAGCAAGAAAGCGATGCCCTTCCCAATCTCTCAATCCTT 780  
QY 824 TTATGCCAGAGAGATCTCAGCTGATGCCAATGTTCCGATGCTGTGAAAGCATGCC 883  
DB 781 TTATGCCAGAGAGATCTCAGCTGATGCCAATGTTCCGATGCTGTGAAAGCATGCC 840  
QY 884 GACGTTCTCTTGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 943  
DB 841 GACGTTCTCTTGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900  
QY 944 CTTTGTGACGCTGTTTGTGTTATGATGATTTTATGCTGATCTTATGTTTAAAG 1003  
DB 901 CTTTGTGACGCTGTTTGTGTTATGATGATTTTATGCTGATCTTATGTTTAAAG 960  
QY 1004 ACAAGTGTAGTGCAATTTATTTATGTCACCTTTCATTAATTAATTAATTAATTA 1063  
DB 961 ACAAGTGTAGTGCAATTTATTTATGTCACCTTTCATTAATTAATTAATTAATTA 1020  
QY 1064 AAAAAAAAAAAAAA 1077  
DB 1021 AAAAAAAAAAAAAA 1034

RESULT 6

AB211212

ID AB211212 standard; cDNA; 562 BP.

AB211212;

20-JAN-2003 (first entry)

Human polynucleotide SEQ ID NO 94.

Human; genome mapping; gene therapy; food supplement; virus; fungus;  
cell-proliferative disorder; neurodegenerative disease; bacterial;  
Parkinson's disease; Alzheimer's disease; autoimmune disease;  
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
arthritis; cytostatic; immunomodulator; neotropic; neuroprotective;  
antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
antiarthritic; gene; ss.

Hom sapiens.

MO200270539-A2.

12-SEP-2002.

05-MAR-2002; 2002MO-US005095.

05-MAR-2001; 2001US-00799451.

(HYSE-) HYSEQ INC.

XX



PI Tang YT, Zhou P, Goodrich RW, Aarndt V, Zhang J, Zhao Q, Ren F;  
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
PI Weinman T, Wang J, Wang D, Drmanac RT;  
XX WPI; 2002-759812/82.  
DR P-PSDB; ABP68995.

XX New polynucleotides comprising sequences assembled from expressed  
PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
PT or coagulation disorders.

PS Claim 1; SEQ ID NO 94; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (1) comprising a  
CC nucleotide sequence selected from any of 948 sequences (AB211119-  
CC AB212066) or their mature protein coding portion, active domain coding  
CC protein or complementary sequences. The polynucleotides are useful for  
CC identifying expressed genes or for physical mapping of human genome. The  
CC encoded polypeptides (ABP68902-ABP68949) are useful as molecular weight  
CC markers, as a food supplement, for generating antibodies, in medical  
CC imaging, screening and diagnostic assays and for treating cell-  
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
CC platelet or coagulation disorders (bacterial, viral, fungal, parasitic),  
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 562 BP; 95 A; 176 C; 174 G; 117 T; 0 U; 0 Other;

Query Match 47.1%; Score 562; DB 6; Length 562;  
Best Local Similarity 100.0%; Pred. No. 1.8e-211;  
Matches 562; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GCGGGAGTGAAGCGCCGAGGAGGCGGCTGACAGACAGAGTACTGACG 75  
DB 1 GCGGGAGTGAAGCGCCGAGGAGGCGGCTGACAGACAGAGTACTGACG 60  
QY 76 GTGGGCTACCTCTGTGAATCACTGACAGAGCTCCCTCAATTTGTGGTGCAGCCCA 135  
DB 61 GTGGGCTACCTCTGTGAATCACTGACAGAGCTCCCTCAATTTGTGGTGCAGCCCA 120  
QY 136 TCAGGAGACCAAGCGGCTGAGAGATGATGAGGATCTTGGCCATGGGGAATGCTGCA 195  
DB 121 TCAGGAGACCAAGCGGCTGAGAGATGATGAGGATCTTGGCCATGGGGAATGCTGCA 180  
QY 196 GGAAGACGACCCCGAGTGAAGCACTACCAAGCCACCAAGAGTATGATTCCTGACA 255  
DB 181 GGAAGACGACCCCGAGTGAAGCACTACCAAGCCACCAAGAGTATGATTCCTGACA 240  
QY 256 GAGCTTCTTCAATAGAGGAGCAATGATGATGATGATGATGATGATGATGATGATGAT 315  
DB 241 GAGCTTCTTCAATAGAGGAGCAATGATGATGATGATGATGATGATGATGATGATGAT 300  
QY 316 GGAAGTGAAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 375  
DB 301 GGAAGTGAAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
QY 376 GAAATGAGGCTTCCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435  
DB 361 GAAATGAGGCTTCCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 436 CCGGCTCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 495  
DB 421 CCGGCTCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
QY 496 CTACCTGTGTCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 555  
DB 481 CTACCTGTGTCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540

QY 556 TGAGAGGAGACTTGCTGGGCTT 577  
DB 541 TGAGAGGAGACTTGCTGGGCTT 562

RESULT 7

ID AAF82505 standard; cDNA; 550 BP.

XX AAF82505;

DT 18-JUN-2001 (first entry)

DE Human breast tumour protein CDNA 15964.

KW Human; breast cancer; tumour; cytostatic; gene therapy; ss.

OS Homo sapiens.

PN WO200127276-A2.

PD 19-APR-2001.

PF 11-OCT-2000; 2000WO-US028255.

PR 12-OCT-1999; 99US-00417031.

PA (CORI-) CORIYA CORP.

PI Harlocker SL, Dillon DC, Xu J;

DR WPI; 2001-273773/28.

PT New polypeptides encoded by polynucleotide sequences over-expressed in  
PT breast tumor tissue are useful to detect, monitor and treat breast  
PT cancer.

PS Claim 13; Page 50; 52pp; English.

CC The present sequence encodes a breast tumour-associated protein. It was  
CC shown to have at least two-fold overexpression in breast tumour tissue.  
CC The invention relates to an isolated polypeptide comprising at least an  
CC immunogenic portion of a breast tumor-specific protein, or its variant  
CC that retains the ability to react with antigen-specific antisera. The  
CC breast tumour polynucleotides, polypeptides and antibodies are useful for  
CC inhibiting development of breast cancer. The polynucleotides may be used  
CC to design primers and probes for detecting and monitoring breast cancer

XX Sequence 550 BP; 102 A; 163 C; 164 G; 118 T; 0 U; 3 Other;

Query Match 31.0%; Score 369; DB 4; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.6e-135;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 CGGCTTGGACACACAGATCTACTGAGGCTGACCTCTGTGAACATCACTGACTGC 106  
DB 1 CGGCTTGGACACACAGATCTACTGAGGCTGACCTCTGTGAACATCACTGACTGC 60  
QY 107 AAGCTCCCTCAATTTCTGTGAGAGCCCATCAGGAGCCACAGAGGCTGGAGGATGATG 166  
DB 61 AAGCTCCCTCAATTTCTGTGAGAGCCCATCAGGAGCCACAGAGGCTGGAGGATGATG 120  
QY 167 CGATCTTGGCAATGAGGAAATGTCAGAGACGACGACCCCGAGTGAAGACCACTAAC 226  
DB 121 CGATCTTGGCAATGAGGAAATGTCAGAGACGACGACCCCGAGTGAAGACCACTAAC 180  
QY 227 CAGGACCAAGAGTATGATTTCTGACAGAGCTTCTCAATAGGGGCAATGATGATGATGAT 286  
DB 181 CAGGACCAAGAGTATGATTTCTGACAGAGCTTCTCAATAGGGGCAATGATGATGATGAT 240  
QY 287 CCAAGGAGGCTTGGAGCCCGCAGACAGAGAGTCCAGAGCTGGTGTGCTCACTGCC 346  
DB 241 CCAAGGAGGCTTGGAGCCCGCAGACAGAGAGTCCAGAGCTGGTGTGCTCACTGCC 300

Qy	34	CCCTTAAATGACCTCAACCGGAGAGCTGGTGAACAATGGGCTTCCGGAGTGGCAATCTGGC	406
Db	301	CCCTTAAATGACCTCAACCGGAGAGCTGGTGAACAATGGGCTTCCGGAGTGGCAATCTGGC	360
Qy	407	AACCATGCT	415
Db	361	AACCATGCT	369

RESULT	
8	
ABX92251	
ID	ABX92251 standard; cDNA; 365 BP.

DT 08-MAY-2003 (first entry)

DE Human ovarian specific nucleic acid DEX0310\_65.

Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer; non-cancerous ovarian disease; gene therapy; vaccine; cytostatic; gynaecological.

OS Homo sapiens.

PN W0200292785-A2

PD 21-NOV-2002.

PF 13-FEB-2002; 2002WO-US022271.

PR 13-FEB-2001; 2001US-0268290P.

XX	
X	C
X	D
X	E
X	F
X	G
X	H
X	I
X	J
X	K
X	L
X	M
X	N
X	O
X	P
X	Q
X	R
X	S
X	T
X	U
X	V
X	W
X	X
X	Y
X	Z

DATE RECEIVED

PI	Sun Y, Lin C;	PI	Sun Y, Lin C;
PI	Sun Y, Lin C;	PI	Sun Y, Lin C;

DR WPI: 2003-120677/11

DR P-ESDB; AB0061003  
XX

PT New isolated OSNA nucleic acid and encoded polypeptide, useful for  
PT identifying, diagnosing, monitoring, staging, imaging and treating  
PT ovarian cancer and non-cancerous diseases in ovarian tissues.

PS Claim 1; Page 185-186; 224pp; English.

CC The invention relates to a new isolated nucleic acid termed ovarian  
CC specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that  
CC encodes any of 53 fully defined protein sequences appearing as AB61018-  
CC AB61070 (termed ovarian specific proteins, OSP); (b) any of 76 fully  
CC defined nucleotide sequences appearing as ABX92187-ABX92652; or (c) a  
CC sequence having at least 60% sequence identity to the nucleic acid  
CC molecule of (a) or (b). Also included are a method for determining the  
CC presence of an ovary specific nucleic acid (OSNA) in a sample, a vector  
CC comprising an OSNA, a host cell comprising the vector, an isolated OSP  
CC polypeptide, an anti-OSP antibody or fragment, a method for determining  
CC the presence of an ovary specific protein in a sample and a vaccine  
CC comprising an OSP or OSNA. The methods and compositions of the present  
CC invention are useful for identifying, diagnosing, monitoring, staging,  
CC imaging and treating ovarian cancer and non-cancerous disease in ovary  
CC tissue. The present sequence is an OSNA of the invention

Sequence 365 BP; 60 A; 117 C; 108 G; 80 T; 0 U; 0 Other;

Query Match	27.3%	Score 326	DB 7	Length 365
Best Local Similarity	100.0%	Pred. No.	1.4e-118	
Matches 326		Conservative	0	Mismatches 0
				Gaps 0

200 GAGGACCCCGAGTGAGGACCACTACCAGCCACCAAGAGGTAGCATTCCTCGACAGAGC 259

Db	40	GACGACCCCCGAGTAGGAGACACTACCCAGGACCCAGAGGTTAGACATTCTCTCGACGAGGC	99
Qy	260	TTCTTCATATAGGAGGCGCATAGTGTCTCCCGAGGAGGATCTCTGCCCCGCGCAGACGAGGCA	31.9
Db	100	TTCTTCATATAGGAGGCGCATAGTGTCTCCCGAGGAGGATCTCTGCCCCGCGCAGACGAGGCA	15.9
Qy	320	GGGCGCAGGCTGGGTGTCTGCTCAGTCCCCCTTCATATAGCTCAACCGGACGTGTGAAC	37.9
Db	160	GGTGCACAGCTGGGTGTCTGCTCAGTCCCCCTTCATATAGCTCAACCGGACGTGTGAAC	21.9
Qy	380	ATGAGCTTTCCGAGTGGCATTTCCGCAACCATGTGTGTGAGCCGGTGAACCTCATCTTG	42.9
Db	220	ATGGGCTTTCCGAGTGGCATTTCCGCAACCATGTGTGTGAGCCGGTGAACCTCATCTTG	27.9
Qy	440	CTGCTCTTCTGTATGATGCTTTGGTGTCTGTGAGCCCTCTCTGTGTGGCTTGTCTAC	49.9
Db	280	CTGCTCTTCTGTATGATGCTTTGGTGTCTGTGAGCCCTCTCTGTGTGGCTTGTCTAC	33.9
Qy	500	CTGGTGTCCAACTGATGACGCGTGG	52.5
Db	340	CTGGTGTCCAACTGATGACGCGTGG	36.5

RESULT 9  
ABX74673/c  
ID ABX74673 standard; cDNA; 517 BP

AC ABX74673;

DT 21-MAR-2003 (first entry)

DE Human cDNA sequence #128 differentially expressed in CC-RCC types.

KW Human; microarray; solid surface; immobilised probe; CC-RCC;

non-aggressive CC-RCC tumour type; clear cell renal carcinoma; KW

XX

[illegible]

X X

C  
C  
C  
C  
C  
C  
C

**X**

12002 - WFM - 67  
XX XX

XX XX  
ZS-PLAN-ZOOT; ZOOTHOO-CZ/HH  
XX

XX (VAND-) VAN HANDEL INDI.  
XX

Haad B, Riodes U, Jakaasul M,  
Fl  
XX

DR WEI; 2003-0406/9/03.

PT New microarray, comp  
DT immobilized to a so

PT New microarray, comprising a matrix of cDNA probe from a set of probes  
PT immobilised to a solid surface in predetermined order, useful in the  
PT prognosis of patients with clear cell renal carcinoma.  
XX  
2S Claim 1; SEQ ID NO 459; 179pp; English.

CC The present invention relates to a micr

The present invention relates to a microarray comprising a matrix of at least one cDNA probe from a set of probes immobilised to a solid surface in a predetermined order, where a row of pixels corresponds to replicates of one distinct probe from the set. The probes are complementary to CC of distinct sequences that are expressed differentially in aggressive as compared to non-aggressive types of clear cell renal carcinoma (CC-RC) and that hybridise to the probes under high stringency conditions. The microarray is useful for the prognosis of patients with CC-RC, wherein aggressive and non-aggressive CC-RC tumour types are characterised by differential expression profiles of genes that hybridise with one or more probes immobilised on the microarray. The arrays are useful for gene expression profiling of tumour and normal tissues. The present sequence represents a human cDNA sequence differentially expressed in aggressive



CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1165 BP; 247 A; 279 C; 342 G; 297 T; 0 U; 0 Other;

Query Match 13.5%; Score 161; DB 5; Length 1165;  
Best Local Similarity 100.0%; Pred. No. 1e-53;  
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 903 GACGAGACTGGGCTTAGGGCAGTGGAAAAATCCAGCTTTTGGACCTGTTT 962  
Db 999 GACGAGACTGGGCTTAGGGCAGTGGAAAAATCCAGCTTTTGGACCTGTTT 1058  
QY 963 GTTTAATGATATATTTTATTTGCTACTTTATTTGTTTGAACAAGTGTAGTGCATTTC 1022  
Db 1059 GTTTAATGATATATTTTATTTGCTACTTTATTTGTTTGAACAAGTGTAGTGCATTTC 1118  
QY 1023 TATTTATTTGACCTTTTCAATAAATAGATTAAAGTAAAA 1063  
Db 1119 TATTTATTTGACCTTTTCAATAAATAGATTAAAGTAAAA 1159

## RESULT 12

AAS64819/c  
ID AAS64819 standard; cDNA; 348 BP.

AC AAS64819;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #623.

Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.

Dmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG00632.

New isolated polynucleotide and encoded polypeptides, useful in

PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

PS Claim 1; SEQ ID NO 623; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
XX coding sequences of the invention. Note: The sequence data for this  
XX patent did not appear in the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences

SEQ Sequence 348 BP; 121 A; 94 C; 58 G; 75 T; 0 U; 0 Other;

Query Match 13.3%; Score 159; DB 5; Length 348;  
Best Local Similarity 100.0%; Pred. No. 7.6e-53;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 904 AGCAGGACTGGGCTTAGGGCAGTGGAAAAATCCAGCTTTTGGACCTGTTT 963  
Db 159 AGCAGGACTGGGCTTAGGGCAGTGGAAAAATCCAGCTTTTGGACCTGTTT 100  
QY 964 TTTAATGATATATTTTATTTGCTACTTTATTTGTTTGAACAAGTGTAGTGCATTCT 1023  
Db 99 TTTAATGATATATTTTATTTGCTACTTTATTTGTTTGAACAAGTGTAGTGCATTCT 40  
QY 1024 ATTATTTTGACCTTTTCAATAAATAGATTAAAGTAAAA 1062  
Db 39 ATTATTTTGACCTTTTCAATAAATAGATTAAAGTAAAA 1

## RESULT 13

AAS64821  
ID AAS64821 standard; cDNA; 587 BP.

AC AAS64821;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #625.

Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US008631.

31-MAR-2000; 2000US-00540217.

23-AUG-2000; 2000US-00649167.

(HYSE-) HYSEQ INC.



XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.  
 XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 XX gene expression in human cervical epithelial cells.  
 XX

PS Claim 25; SEQ ID NO 6156; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes  
 CC (SNP). The present sequence is one such probe. The SNPs are derived  
 CC from human HeLa cells. The SNPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging of  
 CC diseases of the cervix, notably cervical cancer. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 473 BP; 153 A; 80 C; 101 G; 139 T; 0 U; 0 Other;

Query Match 3.2%; Score 38; DB 4; Length 473;

Best Local Similarity 100.0%; Pred. No. 3e-05;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1048 TAGATTTAACTAAAAA 1085

DB 259 TAGATTTAACTAAAAA 296

Search completed: August 22, 2004, 05:00:50  
 Job time : 520 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 22, 2004, 04:34:54 | Search time 2996 Seconds  
(without alignments)

11881.081 Million cell updates/sec

Title: US-10-001-885-31

Sequence: 1192  
1 aaacgcctgagccgcgcg.....gaagagcacacagcagacc 1192

Scoring table: OLIGO NUC  
Gapop 60.0, Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_pig:\*  
27: em\_gss\_vr1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	52.9	708	12	BM675290 UI-E-EJO-
2	586	49.2	688	12	BT521387 603081177
3	510	42.8	931	13	BK455674 BK455674
4	502	42.1	892	13	BQ884216 BQ884216 AGENCOURT

5	491	41.2	714	10	BF568253
6	491	41.2	1104	12	BG311398
7	484	40.6	937	12	BM459665
8	482	40.4	541	14	CA942743
9	480	40.3	980	10	BE904302
10	475	39.8	475	9	AI288193
11	464	38.9	713	12	BI133763
12	455	38.2	502	12	AM952776
13	446	37.4	446	10	BG055477
14	443	37.2	765	13	BK095713
15	431	36.2	533	13	BU733254
16	427	35.8	442	9	AA682907
17	424	35.6	424	9	AA396343
18	423	35.5	423	9	AA947931
19	423	35.5	431	12	BM676116
20	423	35.5	431	12	BM676116
21	423	35.5	431	13	BU741136
22	422	35.4	427	12	BM713794
23	416	34.9	426	12	BM726841
24	411	34.5	440	10	BE179883
25	410	34.4	461	9	AA084096
26	401	33.6	430	10	BF510821
27	395	33.1	420	9	AI197950
28	391	32.8	391	9	AI694330
29	390	32.7	444	14	CB046478
30	386	32.4	386	9	AI968970
31	376	31.5	376	9	AI368631
32	374	31.4	377	9	AM089957
33	369	31.0	369	10	BF591892
34	361	30.3	518	10	AM805228
35	354	29.7	412	9	AI420512
36	352	29.5	505	14	CA942977
37	338	28.4	338	12	BM714109
38	337	28.3	376	9	AA740545
39	321	26.9	435	13	BU738841
40	311	25.1	480	14	H49529
41	301	25.3	425	14	W90083
42	298	25.0	1117	13	BQ891912
43	254	21.3	528	10	AM960624
44	247	20.7	510	14	H49767
45	235	19.7	510	9	AA426015

#### ALIGNMENTS

RESULT 1  
LOCUS BM675290/c 708 bp mRNA linear EST 27-FEB-2002  
DEFINITION UI-E-EJO-ahr-n-23-0-UI.s1 UI-E-EJO Homo sapiens CDNA clone  
VERSION BM675290  
KEYWORDS BM675290.1 GI:18985188  
SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 708)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548

#### COMMENT

Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman

CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 The following repetitive elements were found in this cDNA  
 sequence: 1-42, >AT-richLow\_complexity (matched complement)  
 Seq primer: M13 Forward  
 POLYA=yes.

# FEATURES

source  
 1. 708  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-EJ0-ahr-n-23-0-UI"  
 /tissue\_type="fetal eyes, lens, eye anterior segment,  
 optic nerve, retina, Retina Foveal and Macular, RPE and  
 Choroid"  
 /dev\_stage="fetal and adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-EJ0"  
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-EJ0 is a subcloned cDNA library constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dG)<sub>18</sub> tail. The  
 sequence tags for this library are: fetal eyes,  
 AGATCCAGCA; lens, CGATTAGCA; eye anterior segment,  
 AATGCCGCA; optic nerve, CCATTAGTG; retina, CCGCG;  
 Foveal and Macular, GTCC; RPE and Choroid, ACCCA. This  
 library was created for the program, Gene Discovery in the  
 Visual System, supported by National Eye Institute (NEI).  
 TAG TISSUE=Foveal and Macular Retina  
 TAG LIB=UI-E-EJ0  
 TAG SEO=GTCC"

## ORIGIN

Query Match 52.9%; Score 630; DB 12; Length 708;  
 Best Local Similarity 99.9%; Pred. No. 9,7e-118;  
 Matches 680; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

383 GGGTTTCCGAGGCGCATCTGCGAACCATGCTGTGAGCCGGTACCTCCATCTGCTC 442  
 DB 701 GGGTTTCCGAGGCGCATCTGCGAACCATGCTGTGAGCCGGTACCTCCATCTGCTC 642  
 QY CTCTTCCTGCTCAGTATGCTGTGCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTG 502  
 DB 443 CTCTTCCTGCTCAGTATGCTGTGCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTG 582  
 DB 641 CTCTTCCTGCTCAGTATGCTGTGCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTG 582  
 QY 503 GTGTCCCACTGAGTCAAGCGGTGACCTCTGAGGGCTGATGAGGGGTGCTGTTGTAAGG 562  
 DB 581 GTGTCCCACTGAGTCAAGCGGTGACCTCTGAGGGCTGATGAGGGGTGCTGTTGTAAGG 522  
 QY 563 GACTTGTGGCCCTGTGTTGAGAGAGAGCATTTTGAAGGGGATCTGAGTGGCTTGA 622  
 DB 521 GACTTGTGGCCCTGTGTTGAGAGAGAGCATTTTGAAGGGGATTTGAGTGGCTTGA 462  
 QY 623 AGGTATGATAGAGAGGGGACCAAGGTGTGTTTCCCTTTGTTGTTAAGCGTGAAGCA 682  
 DB 461 AGGTATGATAGAGAGGGGACCAAGGTGTGTTTCCCTTTGTTGTTAAGCGTGAAGCA 402  
 QY 683 GAGGAGACCTGTGCTCAGATTTCCAAAGTGTGGGTGGCTGCTGTTCCCGAGATC 742  
 DB 401 GAGGAGACCTGTGCTCAGATTTCCAAAGTGTGGGTGGCTGCTGTTCCCGAGATC 342  
 QY 743 TTTTAGGTGATAGGGGCTGCAATTAAGTGCAAAATCAGACAAAGAGATGCC 802

DB 341 TTTTAGGTGATAGGGGCTGCAATTAAGTGCAAAATCAGACAAAGAGATGCC 282  
 QY 803 TTCCCAATTTCTCATACTTTTATAGCCGAGAGATCTCAGCTGATGCCAATGTTCC 862  
 DB 281 TTCCCAATTTCTCATACTTTTATAGCCGAGAGATCTCAGCTGATGCCAATGTTCC 222  
 QY 863 GATCCCTGTGGAACAACATGCCAGTCTCTGCTCCCTAGGAGAGAGCACTGGGCTTAG 922  
 DB 221 GATCCCTGTGGAACAACATGCCAGTCTCTGCTCCCTAGGAGAGAGCACTGGGCTTAG 162  
 QY 923 GCAGGTGGAATAATTCAGACTTTTATAGCACTGTTTGTATATAGTATATTTTA 982  
 DB 161 GCAGGTGGAATAATTCAGACTTTTATAGCACTGTTTGTATATAGTATATTTTA 102  
 QY 983 TTGGCTACTTATTTGTTAGACACAGTGTAGTGTGATTTATTTATGACCTTTTCA 1042  
 DB 101 TTGGCTACTTATTTGTTAGACACAGTGTAGTGTGATTTATTTATGACCTTTTCA 42  
 QY 1043 ATTAATGATTTAGTAAAA 1063  
 DB 41 ATTAATGATTTAGTAAAA 21

RESULT 2  
 B1521387 688 bp mRNA linear EST 29-AUG-2001  
 LOCUS 603081177F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5220570 5',  
 DEFINITION mRNA sequence.  
 B1521387  
 VERSION B1521387.1 GI:15346179  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 1 (bases 1 to 688)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.lnl.gov>  
 Plate: LAM11554 row: 1 column: 19  
 High quality sequence, STOP: 688.

## FEATURES

### source

1. 688  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5220570"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_120"  
 /note="Organ: pooled pancreas and spleen; Vector:  
 PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 kb, insert size range 1-2.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 49.2%; Score 586; DB 12; Length 688;  
 Best Local Similarity 99.7%; Pred. No. 6.6e-109;  
 Matches 686; Conservative 0; Mismatches 2; Indels 0; Gaps 0;





SOURCE	ORGANISM	Homosapiens (human)
REFERENCE	Authors	Human sapiens
REFERENCE	Title	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	Journal	1 (bases 1 to 892)
REFERENCE	Journal	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
REFERENCE	Journal	National Institutes of Health, Mammalian Gene Collection (MGC)
REFERENCE	Journal	Unpublished (1999)
REFERENCE	Journal	Contact: Robert Strausberg, Ph.D.
REFERENCE	Journal	Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a>
REFERENCE	Journal	Tissue Procurement: Dr. James R. Lupski
REFERENCE	Journal	CDNA Library Preparation: Life Technologies, Inc.
REFERENCE	Journal	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
REFERENCE	Journal	DNA Sequencing by: Agencourt Bioscience Corporation
REFERENCE	Journal	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LML at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>
REFERENCE	Journal	Plate: LML13607 row: 1 column: 12
REFERENCE	Journal	High quality sequence slope: 497.
REFERENCE	Journal	Location/Qualifiers
REFERENCE	Journal	1..892
REFERENCE	Journal	/organism="Homo sapiens"
REFERENCE	Journal	/mol_type="mRNA"
REFERENCE	Journal	/db_xref="taxon:9606"
REFERENCE	Journal	/clone="IMAGE:6197459"
REFERENCE	Journal	/sex="male"
REFERENCE	Journal	/tissue_type="sciatic nerve"
REFERENCE	Journal	/dev_stage="adult, 70 yr"
REFERENCE	Journal	/lab_host="DH10B"
REFERENCE	Journal	/clone_lib="Lupski sciatic nerve"
REFERENCE	Journal	/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; CDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TGACCCCAAGCGGCTCG-3' and 5'-GACTGTTGATGATCGGAGCGGCGCTT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."
ORIGIN		
Query Match	42.1%; Score 502; DB 13; Length 892;	
Best Local Similarity	99.7%; Pred. No. 3.6e-92;	
Matches 602; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
Db	292 GGGTCTCTGACCCCGCCAGACAGGACAGGCGCCAGGCTGGATGCTGCTCAATGCCCTT	351
Db	1 GGGTCTCTGACCCCGCCAGACAGGACAGGCGCGGTCAGGCTGGATGCTGCTCAATGCCCTT	60
Db	352 CAATGACCTCAACCGGACAGCTGTGAAATATGGGCTTTCCGAGTGGCATTTCGGCAACA	411
Db	61 CAATGACCTCAACCGGACAGCTGTGAAATATGGGCTTTCCGAGTGGCATTTCGGCAACA	120
Db	412 TGTGTGAGAGCCGGTGAACCTGCATCTCTCTCTTCTGTCTGATGATGCTTGTGTTGG	471
Db	121 TGTGTGAGAGCCGGTGAACCTGCATCTCTCTCTTCTGTCTGATGATGCTTGTGTTGG	180
Db	472 TGTGTCTCTCTGTGTGGCTTTGTCTTACCTGTGTCCCACTGATGAGCGGTGACCTCT	531
Db	181 TGTGTCTCTCTGTGTGGCTTTGTCTTACCTGTGTCCCACTGATGAGCGGTGACCTCT	240
Db	532 GAGGCGTGAATAGGGGGGGGTTTGTGAAGGAGACTTGTGAGGCTTGTGTGAGACAG	591
Db	241 GAGGCGTGAATAGGGGGGGGTTTGTGAAGGAGACTTGTGAGGCTTGTGTGAGACAG	300
Db	592 CATATTGAGAGGGGATCTGTGTGTCCTTGAAGGTATGATCAGAGAGGACCAAGGTG	651
Db	301 CATATTGAGAGGGGATCTGTGTGTCCTTGAAGGTATGATCAGAGAGGACCAAGGTG	360
Db	652 TGTGTTTCCCTTTGTGTTAAGCGTGAAGGAGAGGAGAGGTTAAGTCAAGCATTTCCAA	711
Db	361 TGTGTTTCCCTTTGTGTTAAGCGTGAAGGAGAGGAGAGGTTAAGTCAAGCATTTCCAA	420

QY	712	GTGTGGGTGGTGGTCGGTTGGTTCTCCGAGATACCTTTAGTGGTATGGGGCTGCATTAGT	772
Db	421	GTGTGGGTGGTGGTCGGTTGGTTCTCCGAGATACCTTTAGTGGTATGGGGCTGCATTAGT	480
QY	772	GGCACAATAATCAGACGCAAGAGCGATGCCCTTCCCAATTCTCTCAATCTTTATGCCG	831
Db	481	GGCACAATAATCAGACGCAAGAGCGATGCCCTTCCCAATTCTCTCAATCTTTATGCCG	540
QY	832	AGAAGATCTCAGCTGGATGCGCAACATTTCCGATGCGCTGTGAAGACATGCCGCGCTCC	891
Db	541	AGAAGATCTCAGCTGGATGCGCAACATTTCCGATGCGCTGTGAAGACATGCCGCGCTCC	600
QY	892	CTCT 895	
Db	601	CTCT 604	
RESULT 5			
BF568253			
LOCUS			
DEFINITION	BF568253	714 bp	mRNA linear EST 12-DEC-2000
ACCESSION	BF568253		
VERSION	BF568253.1		
KEYWORDS	EST		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Ekikayote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 714)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaps-remail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: ILCM159 row: j column: 07		
	High quality sequence stop: 712.		
FEATURES			
source			
	1..714		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4300518"		
	/tissue_type="epithelioid carcinoma cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_id="NIH MGC 42"		
	/notes="Organ: pancreas; Vector: pOTB7; Site: 1; XhoI;		
	Site 2; EcoRI; CDNA made by oligo-dt priming.		
	Directionally cloned into EcoRI/XhoI sites using the		
	following 5' adaptor: GGCAAGAG(G). Size-selected 500bp		
	for average insert size 1.8kb. Library constructed by Ling		
	Hong in the laboratory of Gerald M. Rubin (University of		
	California, Berkeley) using ZAP-CDNA synthesis kit		
	(Stratagene) and Superscript II RT (Life Technologies).		
	Note: this is a NIH-MGC Library.  "		
ORIGIN			
Query Match	41.2%;	Score 491;	DB 10; Length 714;
Best Local Similarity	99.8%;	Pred. No. 7.2e-90;	
Matches 541; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
10	GGAGCCCCCGGAGTGCAGCCCGCGAGGCTCGGCTTCGACAGACAGCATCTTA	69	
2	GGAGCCCCCGGAGTGCAGCCCGCGAGGCTCGGCTTCGACAGACAGCATCTTA	61	

QY 70 CTGAGGCTGGGTCACCTCTGTGTGATCACTGATCGCAAGCCCTCCCTCAATTTCTGTGTC 129  
Db 62 CTCAGGTGGTGCACCTCTGTGTGATCACTGATCGCAAGCCCTCCCTCAATTTCTGTGTC 121  
QY 130 AGCCCATCGAGGACCCCAAGCCCTGGAGAGATGTGCGGATCTTGGCCCATGGGAAAT 189  
Db 122 AGCCCATCGAGGACCCCAAGCCCTGGAGAGATGTGCGGATCTTGGCCCATGGGAAAT 181  
QY 190 CGTGAGGAG 249  
Db 182 CGTGAGGAG 241  
QY 250 TCGACAGAGCTTCTTCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 309  
Db 242 TCGACAGAGCTTCTTCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301  
QY 310 GAG 369  
Db 302 GAG 361  
QY 370 GGTGTGTAACATGAGGCTTCCGCAAGTGCATCTCGGCAACATGTGTGTGAGCCGGTAC 429  
Db 362 GGTGTGTAACATGAGGCTTCCGCAAGTGCATCTCGGCAACATGTGTGTGAGCCGGTAC 421  
QY 430 CTCATCTCTGCT 489  
Db 422 CTCATCTCTGCT 481  
QY 490 CTTGTCTACCTGT 549  
Db 482 CTTGTCTACCTGT 541  
QY 550 GT 551  
Db 542 GT 543

RESULT 6  
BG31398 1104 bp mRNA linear EST 27-FEB-2001  
LOCUS 602433369P1 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:4550896 5',  
DEFINITION mRNA sequence.  
BG31398  
ACCESSION BG31398.1 GI:13137937  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1104)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: DCTD/DMP/Gazdar  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: L10M1244 row: j column: 17  
High quality sequence stop: 636.  
Location/Qualifiers  
1. 1104  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_image="4550896"  
/tissue\_type="large cell carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_18"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source

/note="Organ: lung; Vector: pOT7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
Laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC library."

Query Match 41.2%; Score 491; DB 12; Length 1104;  
Best Local Similarity 99.8%; Pred. No. 4,7e-90;  
Matches 541; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 GGAAGCCGCGGAGATGAGAGCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 69  
Db 1 GGAAGCCGCGGAGATGAGAGCCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60  
QY 70 CTCAGCGTGGGTACCTCTGTGAACATCACTGACTGACCTGCAATTTCTGTGTC 129  
Db 61 CTCAGCGTGGGTACCTCTGTGAACATCACTGACTGACCTGCAATTTCTGTGTC 120  
QY 130 AGCCCATCGAGGACCCCAAGCCCTGGAGAGATGTGCGGATCTTGGCCCATGGGAAAT 189  
Db 121 AGCCCATCGAGGACCCCAAGCCCTGGAGAGATGTGCGGATCTTGGCCCATGGGAAAT 180  
QY 190 CGTGAGGAG 249  
Db 181 CGTGAGGAG 240  
QY 250 TCGACAGAGCTTCTTCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 309  
Db 241 TCGACAGAGCTTCTTCATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
QY 310 GAG 369  
Db 301 GAG 360  
QY 370 GGTGTGTAACATGAGGCTTCCGCAAGTGCATCTCGGCAACATGTGTGTGAGCCGGTAC 429  
Db 361 GGTGTGTAACATGAGGCTTCCGCAAGTGCATCTCGGCAACATGTGTGTGAGCCGGTAC 420  
QY 430 CTCATCTCTGCT 489  
Db 421 CTCATCTCTGCT 480  
QY 490 CTTGTCTACCTGT 549  
Db 481 CTTGTCTACCTGT 540  
QY 550 GT 551  
Db 541 GT 542

RESULT 7  
BM459665 937 bp mRNA linear EST 05-FEB-2002  
LOCUS AGENCOURL 6417963 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5534295  
DEFINITION 5', mRNA sequence.  
BM459665  
ACCESSION BM459665.1 GI:18508705  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 937)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Tissue Procurement: ATCC  
 cDNA library Preparation: Life Technologies, Inc.  
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LIML at:  
<http://image.jml.gov>  
 Place: LHAM1220 row: 1 column: 16  
 High quality sequence stop: 585.

## FEATURES

Location/Qualifiers  
 1..937  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5534295"  
 /tissue\_type="leiomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2.1 kb."

## ORIGIN

Query Match 40.6%; Score 484; DB 12; Length 937;  
 Best Local Similarity 99.5%; Pred. No. 1.4e-88;  
 Matches 634; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 328 GCTGGGTCCTGCTCACTCCCTCTTCAATACCTCAACCGGAGCTGGTGAACATGGGCTT 387
DB 1 GCTGGGTCCTGCTCACTCCCTCTTCAATACCTCAACCGGAGCTGGTGAACATGGGCTT 60
QY 388 TCCGAGTGGAGCTTCGCGCAACCATCTGTGAGCGCGGTGACCTCCATCTGCTCTT 447
DB 61 TCCGAGTGGAGCTTCGCGCAACCATCTGTGAGCGCGGTGACCTCCATCTGCTCTT 120
QY 448 CCTGCTCATGATGCTTGTGTTGTTGTTGCTCTCTCTGTTGGCTTGTCTACCTGTTGTC 507
DB 121 CCTGCTCATGATGCTTGTGTTGTTGTTGCTCTCTCTGTTGGCTTGTCTACCTGTTGTC 180
QY 508 CCACCTGAGTACGCGGTGACCTCTGAGGCTGATAGGGGTGGTGTGTTGAAGGACTT 567
DB 181 CCACCTGAGTACGCGGTGACCTCTGAGGCTGATAGGGGTGGTGTGTTGAAGGACTT 240
QY 568 GCTGGGCTTGTGTGAGAGCAGGCAATTTGAGAGGGGATCTGTGTGCTTGAAGTGA 627
DB 241 GCTGGGCTTGTGTGAGAGCAGGCAATTTGAGAGGGGATCTGTGTGCTTGAAGTGA 300
QY 628 TGTATGAGAGGGGACACAGGTGTGTGTTCCCTTGTGTGTTAGCGTGAAGGAGG 687
DB 301 TGTATGAGAGGGGACACAGGTGTGTGTTCCCTTGTGTGTTAGCGTGAAGGAGG 360
QY 688 AGACGTTAGTCCAGCATTTCCAAAGTGTGGTGGGTCCGTTGGTCCCGAGTACTTTTA 747
DB 361 AGACGTTAGTCCAGCATTTCCAAAGTGTGGTGGGTCCGTTGGTCCCGAGTACTTTTA 420
QY 748 GGTGTATGGGGCTGCTGATTAAGTGGCAAAATTCAGACCAAGAAAGCATGCCCTTCCC 807
DB 421 GGTGTATGGGGCTGCTGATTAAGTGGCAAAATTCAGACCAAGAAAGCATGCCCTTCCC 480
QY 808 AATTCCTCAATCTTTTATGCGGAGAAATCTCAGCTGATGCCAATGTTCCGATGC 867
DB 481 AATTCCTCAATCTTTTATGCGGAGAAATCTCAGCTGATGCCAATGTTCCGATGC 540
QY 868 CTGTGAGAGCATGCGGAGCTCTCTCTGTGCTAGGAGACGACTTGGGCTTAGGGCAG 927
DB 541 CTGTGAGAGCATGCGGAGCTCTCTCTGTGCTAGGAGACGACTTGGGCTTAGGGCAG 600
QY 928 TGAATAAATTCAGACTTTTATAGCACTGTTTGT 964
DB 601 TGAATAAATTCAGACTTTTATAGCACTGTTTGT 637

```

CA942743/c  
 LOCUS CA942743 541 bp mRNA linear EST 30-DEC-2002  
 DEFINITION ir66f05.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6607449 3',  
 mRNA sequence.  
 CA942743  
 CA942743.1 GI:27431223  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 541)  
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
 Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistein, A.,  
 Schmitt, A., Heising, B., Ritzer, E., Ronko, I., Bennett, J.,  
 Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagarashvili, R.,  
 Williams, T., Jackson, Y., and Bowers, Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)  
 Other ESTs: ir66f05.Y1  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohpc.harvard.edu  
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
 Washington University Genome Sequencing Center For information on  
 obtaining a clone please contact: Dr. Hiroshi Inoue  
 (hinoue@im.wustl.edu)  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 455.  
 Location/Qualifiers  
 1..541  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6607449"  
 /tissue\_type="Purified pancreatic islet"  
 /lab\_host="DH10B"  
 /clone\_1ib="HR85 islet"  
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1:  
 NotI; Site 2: XhoI; cDNA made by oligo-dT priming.  
 Size-selected on agarose gel. Average insert size ~1kb. 5'  
 XhoI site was destroyed after directional cloning.  
 Amplified once. Contact information: Hiroshi Inoue, MD,  
 Metabolism Div. (Alan Permutt Lab), Washington University  
 School of Medicine, Box 8127, 660 South Euclid Ave., St.  
 Louis, MO 63110, E-mail: hinoue@im.wustl.edu, Tel:  
 314-362-1916, Fax: 314-747-2692."

## FEATURES

## source

## ORIGIN

Query Match 40.4%; Score 482; DB 14; Length 541;  
 Best Local Similarity 99.8%; Pred. No. 5.9e-88;  
 Matches 532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 531 TGAAGGCTGATAGGGGTGGGTTGTTGAGAGGACTTCTGGGCTTGTGTAGAGCAG 590
DB 541 TGAAGGCTGATAGGGGTGGGTTGTTGAGAGGACTTCTGGGCTTGTGTAGAGCAG 482
QY 591 GCATATTTGAGGGGAGCTGTGTGTGCTTGAAGGTATGATCAGAGGGGACCAAGGT 650
DB 481 GCATATTTGAGGGGAGCTGTGTGTGCTTGAAGGTATGATCAGAGGGGACCAAGGT 422
QY 651 GTGTGTTCCCTTTGTGTATAGCGTGAAGGAGAGGAGCGATTAGTCCAGCATTTCCAA 710
DB 421 GTGTGTTCCCTTTGTGTATAGCGTGAAGGAGAGGAGCGATTAGTCCAGCATTTCCAA 362
QY 711 AGTGTGGTGGTCCCTTGTGTCCGAGTACTTTTATAGTGTGATGGGCGCTCATTAAG 770

```

Db 361 AGTGTGGGTGGTCCGTTGGTTCCAGATCTTTAGGTGTATGGGGCCCTGCATTAG 302  
QY 771 TGGCACAAAATCAGAGCAAGAAAGCATGCCCTTCCCAATTCCTCAATCCTTTATGCC 830  
Db 301 TGGCACAAAATCAGAGCAAGAAAGCATGCCCTTCCCAATTCCTCAATCCTTTATGCC 242  
QY 831 GAGAAAGATCTCAGCTGGAATGCCCAATTCCTCAATTCCTGGAAGCATGCCAGCTCT 890  
Db 241 GAGAAAGATCTCAGCTGGAATGCCCAATTCCTCAATTCCTGGAAGCATGCCAGCTCT 182  
QY 891 CCTGTGCTTAGGAGAGCAAGATCTGGGCTTAGGAGAGAGGAAAAATTCAGACTTTT 950  
Db 181 CCTGTGCTTAGGAGAGCAAGATCTGGGCTTAGGAGAGAGGAAAAATTCAGACTTTT 122  
QY 951 AGCACTGTTTGTATATGATATATTTTATTTGCTACTTTATTTGTTTGGCAAGT 1010  
Db 121 AGCACTGTTTGTATATGATATATTTTATTTGCTACTTTATTTGTTTGGCAAGT 62  
QY 1011 GTAGTGGGATCTATTTATTTGACCTTTTCATTAATAGATTAGTAAAAA 1063  
Db 61 GTAGTGGGATCTATTTATTTGACCTTTTCATTAATAGATTAGTAAAAA 9

RESULT 9  
BE904302 980 bp mRNA linear EST 20-OCT-2000  
LOCUS 601494571P2 NIH\_MGC\_70 Homo sapiens CDNA clone IMAGE:389656 5',  
DEFINITION mRNA sequence.  
ACCESSION BE904302 GI:10396415  
VERSION BE904302.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 980)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Inocyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: L1AM9689 row: n column: 17  
High quality sequence stop: 711.  
Location/Qualifiers  
1..980  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:389656"  
/tissue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1lb="NIH MGC 70"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI. Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."

FEATURES  
source  
1..980  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:389656"  
/tissue\_type="epithelioid carcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1lb="NIH MGC 70"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI. Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."

ORIGIN  
Query Match 40.3%; Score 480; DB 10; Length 980;  
Best Local Similarity 99.8%; Pred. No. 8,5e-88;  
Matches 530; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 22 AGTGAAGCCGCGAGGCGCGAGTGCCTGAGACACAGATCTACTCAGCGGTG 81  
Db 1 AGTGAAGCCGCGAGGCGCGAGTGCCTGAGACACAGATCTACTCAGCGGTG 60

QY 82 CACCTGTGAACATCACTGACAGAGCCCTCCATTTCTGTGACGCCCATCAGG 141  
Db 61 CACCTGTGAACATCACTGACAGAGCCCTCCATTTCTGTGACGCCCATCAGG 120  
QY 142 ACCCAGAGGCGCTGGAGAGATGTGCGGATCTTTGGCAATGGGGAATGTGAGAGCA 201  
Db 121 ACCCAGAGGCGCTGGAGAGATGTGCGGATCTTTGGCAATGGGGAATGTGAGAGCA 180  
QY 202 CGACCCCGAGTGAAGACCACTACCCAGCAACCAAGAGTAGATCTCTGACAGACT 261  
Db 181 CGACCCCGAGTGAAGACCACTACCCAGCAACCAAGAGTAGATCTCTGACAGACT 240  
QY 262 CTTCATAGAGGCGCATGTGCTCCCAAGGGGCTCTGCCCCCGAGAGAGAGG 321  
Db 241 CTTCATAGAGGCGCATGTGCTCCCAAGGGGCTCTGCCCCCGAGAGAGAGG 300  
QY 322 TGGCAGGCTGGGTGCTGCTGAGTCCCTTCATAGACCTCAACCGAGCTGTGAACAT 381  
Db 301 TGGCAGGCTGGGTGCTGCTGAGTCCCTTCATAGACCTCAACCGAGCTGTGAACAT 360  
QY 382 GGGCTTTCGAGTGGCATCTGAGCAACATGTGTGAGACCGGCTGACTCTGCT 441  
Db 361 GGGCTTTCGAGTGGCATCTGAGCAACATGTGTGAGACCGGCTGACTCTGCT 420  
QY 442 CTTCTTCTCTCATATGCTGTGTGTGCTGAGCTCTGCTGAGCTGTGACT 501  
Db 421 CTTCTTCTCTCATATGCTGTGTGTGCTGAGCTCTGCTGAGCTGTGACT 480  
QY 502 GGTGTCCACCTGAGTGAAGGCTGAGGCTGATGAGGCTGAGGCT 552  
Db 481 GGTGTCCACCTGAGTGAAGGCTGAGGCTGATGAGGCTGAGGCT 531

RESULT 10  
A1288193 475 bp mRNA linear EST 29-JAN-1999  
LOCUS A1288193/c  
DEFINITION c190e03.x1 Soares NHPu S1 Homo sapiens CDNA clone IMAGE:1879612  
3, similar to contains MER22.tl MER22 repetitive element ;, mRNA  
sequence.  
ACCESSION A1288193  
VERSION A1288193  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 475)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 1129 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 409.  
Location/Qualifiers  
1..475  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1879612"  
/tissue\_type="pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/clone\_1lb="Soares NHPu S1"  
/note="Organ: mixed (see below); Vector: pT773D-Pac  
(Pharmacia) with a modified polylinker; Site 1: Not I;  
Site 2: Eco RI. Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NHP, pregnant uterus  
NHPu, and fetal heart NHPu19W) were mixed, and ss circles

FEATURES  
source  
1..475  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1879612"  
/tissue\_type="pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/clone\_1lb="Soares NHPu S1"  
/note="Organ: mixed (see below); Vector: pT773D-Pac  
(Pharmacia) with a modified polylinker; Site 1: Not I;  
Site 2: Eco RI. Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NHP, pregnant uterus  
NHPu, and fetal heart NHPu19W) were mixed, and ss circles

were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

## ORIGIN

Query Match 39.8%; Score 475; DB 9; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.7e-86;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 588 CAGCATATTTGGAGGGGATCTGGTGTGCTTGAAGATGATCAAGAGGGGACCA 647  
DB 475 CAGCATATTTGGAGGGGATCTGGTGTGCTTGAAGATGATCAAGAGGGGACCA 416  
QY 648 GGTGTGTGTTCCCTTGTGTTAAGCGTGAAGAGGAGAGCTTAGTCCAGCATTT 707  
DB 415 GGTGTGTGTTCCCTTGTGTTAAGCGTGAAGAGGAGAGCTTAGTCCAGCATTT 356  
QY 708 CAAAGTGTGGTGGGTCCTGTTGTCGAGATCTTTAAGTGTATGGGGCTGCAT 767  
DB 355 CAAAGTGTGGTGGGTCCTGTTGTCGAGATCTTTAAGTGTATGGGGCTGCAT 296  
QY 768 AAGTGGCAAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 827  
DB 295 AAGTGGCAAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 236  
QY 828 GCCAGAGAGATCTCAAGTGTGATCCAAATGTTCCGATGCTGTGAGAGAGAG 887  
DB 235 GCCAGAGAGATCTCAAGTGTGATCCAAATGTTCCGATGCTGTGAGAGAG 176  
QY 888 TCTCCTCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 947  
DB 175 TCTCCTCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 116  
QY 948 TTAGACATGTTTTTTTAAATGATATTTTATTTGCTACTTATGTTAGACA 1007  
DB 115 TTAGACATGTTTTTTTAAATGATATTTTATTTGCTACTTATGTTAGACA 56  
QY 1008 GTGTGTGTGCTTCAATTTATTTGACCTTTCAATAAATGATTTAAGTAAA 1062  
DB 55 GTGTGTGTGCTTCAATTTATTTGACCTTTCAATAAATGATTTAAGTAAA 1

## RESULT 11

BI193763 713 bp mRNA linear EST 10-JUL-2001  
LOCUS BI193763  
DEFINITION 602348418F1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:5091412 5',  
mRNA sequence.  
ACCESSION BI193763  
VERSION BI193763  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 713)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L10C1855 row: d column: 05  
High quality sequence stop 709

## source

1..713  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5091412"  
/tissue\_type="epithelioid carcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_id="NIH\_MGC\_42"  
/note="Organ: pancreas; Vector: pOT7; Site: 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACAGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 38.9%; Score 464; DB 12; Length 713;  
Best Local Similarity 99.8%; Pred. No. 1.9e-84;  
Matches 514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 142 ACCCAAGCGCTTGGAGAGATGTCGAGATCTTGGCAATGGGGAATCTGCAAGCA 201  
DB 12 ACCCAAGCGCTTGGAGAGATGTCGAGATCTTGGCAATGGGGAATCTGCAAGCA 71  
QY 202 CGACCCCGAGTGAAGCACTACCAAGCAAGAGATGAGATTCCTGACAGAGCTT 261  
DB 72 CGACCCCGAGTGAAGCACTACCAAGCAAGAGATGAGATTCCTGACAGAGCTT 131  
QY 262 CTTCAATGAGGAGCCATGCTCTCCCAAGGAGGTCCTGAGCCCGCAGAGAGCA 321  
DB 132 CTTCAATGAGGAGCCATGCTCTCCCAAGGAGGTCCTGAGCCCGCAGAGAGCA 191  
QY 322 TGCCAGGCTGGTGTCTAGTCCCTTCAATGACCTCAACCGGAGAGTGTGAAC 381  
DB 192 TGCCAGGCTGGTGTCTAGTCCCTTCAATGACCTCAACCGGAGAGTGTGAAC 251  
QY 382 GAGCTTTCGCAAGTGCATCTCGCAACATGCTGTGAGCCGAGTCACTCTGCT 441  
DB 252 GAGCTTTCGCAAGTGCATCTCGCAACATGCTGTGAGCCGAGTCACTCTGCT 311  
QY 442 CTTCTTCTCTCATATGTTGTGTGTGCTGTGCTCTCTGCTGTGCTGTGCT 501  
DB 312 CTTCTTCTCTCATATGTTGTGTGTGCTGTGCTCTCTGCTGTGCTGTGCT 371  
QY 502 GGTGTCCACCTGAGTCAAGCGGTGACCTTGAAGGCTGATGAGGCTGTGAGAG 561  
DB 372 GGTGTCCACCTGAGTCAAGCGGTGACCTTGAAGGCTGATGAGGCTGTGAGAG 431  
QY 562 GACTTGTCTGGGCTTGTGTGTGAGAGAGAGAGATTTGAGAGGAGTGTGCT 621  
DB 432 GACTTGTCTGGGCTTGTGTGTGAGAGAGAGAGATTTGAGAGGAGTGTGCT 491  
QY 622 AAGTATGATCAGAGAGGAGACCAAGTGTGTGT 656  
DB 492 AAGTATGATCAGAGAGGAGACCAAGTGTGTGT 526

## RESULT 12

AM952776 502 bp mRNA linear EST 01-JUN-2000  
LOCUS AM952776  
DEFINITION EST164846 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.  
ACCESSION AM952776  
VERSION AM952776.1 GI:8142459  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 502)  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE  
Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.  
JOURNAL  
Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray  
COMMENT  
Unpublished (2000)  
Contact: John Quackenbush  
The Institute for Genomic Research  
7112 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johng@tigr.org  
Plate: 43  
Seq primer: Reverse.

FEATURES  
Source  
Location/Qualifiers  
1..502  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE resequences, YAGB"  
/note="Vector: pBluescriptSKm"

#### ORIGIN

Query Match 38.2%; Score 455; DB 10; Length 502;  
Best Local Similarity 100.0%; Pred. No. 1.7e-82;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 ACATGGGCTTCCGAGTGGCATCTGGCAACCATCTGTGAGCCGGTGAACCTCCATCC 437  
DB 1 ACATGGGCTTCCGAGTGGCATCTGGCAACCATCTGTGAGCCGGTGAACCTCCATCC 60  
QY 438 TGGCTCTCTCCGCTCATGATGCTGTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCT 497  
DB 61 TGGCTCTCTCCGCTCATGATGCTGTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCT 120  
QY 498 ACTGCTGCTCCAGCTGAGTCAAGCGGTGACCTCTGAGGCTGATAGGGGTGGCTTTG 557  
DB 121 ACTGCTGCTCCAGCTGAGTCAAGCGGTGACCTCTGAGGCTGATAGGGGTGGCTTTG 180  
QY 558 AGAGGAGCTGCTGAGGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 617  
DB 181 AGAGGAGCTGCTGAGGCTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
QY 618 CTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 677  
DB 241 CTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
QY 678 AGGCAAG 737  
DB 301 AGGCAAG 360  
QY 738 GATACCTTTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 797  
DB 361 GATACCTTTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420  
QY 798 TGGCTCTCCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832  
DB 421 TGGCTCTCCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 455

RESULT 13  
BG055477/c 446 bp mRNA linear EST 25-JAN-2001  
LOCUS nad5sh01.x1 NCI\_CGAP\_Lu24 Homo sapiens cDNA clone IMAGE:3407472 3',  
DEFINITION mRNA sequence.  
ACCESSION BG055477  
VERSION BG055477.1 GI:12516626  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
REFERENCE 1 (bases 1 to 446)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLT, send email to:  
info@image.lnlt.gov  
Seq primer: -400P from Gibeo  
High quality sequence stop: 438.

FEATURES  
Source  
Location/Qualifiers  
1..446  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:3407472"  
/issue\_type="carcinoid"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Lu24"  
/note="Organ: lung; Vector: pUT3D-Pac (Pharmacia) with a  
modified polylinker; Plasmid DNA from the normalized  
library NCI-CGAP Lu5 was prepared, and ss circles were  
made in vitro. Following HAP purification, this DNA was  
used as tracer in a subtractive hybridization reaction.  
The driver was PCR-amplified cDNAs from a pool of 5,000  
clones made from the same library (clonids  
141920-141991 and 1520904-1522439). Subtraction by Bento  
Soares and M. Fatima Bonaldo."

#### ORIGIN

Query Match 37.4%; Score 446; DB 10; Length 446;  
Best Local Similarity 100.0%; Pred. No. 1.2e-80;  
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 588 CAGGCAATTTGGAGGGAGATGCTGCTGCTTGAAGTATGATCAGAGAGAGAGAGAGAG 647  
DB 446 CAGGCAATTTGGAGGGAGATGCTGCTGCTTGAAGTATGATCAGAGAGAGAGAGAGAG 387  
QY 648 GGTGTGCTTCCCTTTGTTAGGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 707  
DB 386 GGTGTGCTTCCCTTTGTTAGGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 327  
QY 708 CAAGGTGAGTGGGTCCTGTTGTTCCGAGATACCTTTAGGTGATGAGGAGAGAGAG 767  
DB 326 CAAGGTGAGTGGGTCCTGTTGTTCCGAGATACCTTTAGGTGATGAGGAGAGAGAG 267  
QY 768 AAGTGGCAAAATCAG 827  
DB 266 AAGTGGCAAAATCAG 207  
QY 828 GCCGAGAGAGATCTAGCTGATGAGCAATGTTCCGATGCTGTGAGAGAGAGAGAGAG 887  
DB 206 GCCGAGAGAGATCTAGCTGATGAGCAATGTTCCGATGCTGTGAGAGAGAGAGAGAG 147  
QY 888 TCTCTCTGCTTAGGAG 947  
DB 146 TCTCTCTGCTTAGGAG 87  
QY 948 TTAGCACTGTTTGTGTTTATGATATTTTATGATGATGATGATGATGATGATGATG 1007  
DB 86 TTAGCACTGTTTGTGTTTATGATATTTTATGATGATGATGATGATGATGATGATG 27  
QY 1008 GTGGTATGAGATCTATTTATTTATG 1033  
DB 26 GTGGTATGAGATCTATTTATTTATG 1

RESULT 14

	LOCUS	BX095713	765 bp	RNA	linear	EST 04-FEB-2003
	DEFINITION	BX095713 Soares_total fetus NB2HF8_9w Homo sapiens cDNA clone				
	ACCESSION	BX095713				
	VERSION	BX095713.1 GI:27842493				
	KEYWORDS	EST.				
	SOURCE	Homo sapiens (human)				
	ORGANISM	Homo sapiens				
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
	AUTHORS	1 (bases 1 to 765) Bert,L., Hell,O., Hennig,S., Neubert,P., Partsch,E., Peters,M., Radejof,U., Schneider,D. and Korn,B.				
	TITLE	Human Unigeneset - RZPD				
	JOURNAL	Unpublished (2003)				
	COMMENT	Contact: Ina Rolfs RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany RZPDLIB; I.M.A.G.E. CDNA Clone Collection; RZPDLIB; I.M.A.G.E. CDNA Clone Collection; Human Unigeneset - RZPD (RZPDLIB No.972) <a href="http://www.rzpd.de/Clovecards/cgi-bin/showLib.pl.cgi?response2libNo=972">http://www.rzpd.de/Clovecards/cgi-bin/showLib.pl.cgi?response2libNo=972</a> Contact: Ina Rolfs RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Tel: +49 30 32639 101 Fax: +49 30 32639 111 <a href="http://www.rzpd.de">www.rzpd.de</a> This clone is available royalty-free from RZPD; contact RZPD ( <a href="mailto:clone@rzpd.de">clone@rzpd.de</a> ) for further information. Seq primer: M3T, Primer sequence: TTTCACACAGAAACGCTATCAC. Location/Qualifiers				
FEATURES	source	1..765				
		/organism="Homo sapiens"				
		/mol_type="mRNA"				
		/db_xref="taxon:9606"				
		/clone="IMAGP98B081942 ; IMAGE:787879"				
		/dev_stage="8-9 weeks"				
		/lab_host="DH10B"				
		/clone_lib="Soares total_fetus NB2HF8_9w"				
		/note="Vector: pRTT3-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - oligo(dT) primer [5' TGTAACCATCTGAAGTGAGCGGCCGCCTTAATTTTT TTTT TTTT 3'] . Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."				
ORIGIN						
	Query Match	37.2%	Score	443;	DB 13;	Length 765;
	Best Local Similarity	99.5%;	Pred.	No. 2.8e-80;		
	Matches 593;	Conservative	0;	Mismatches	3;	Indels 0; Gaps 0;
Dy	117	CAATTCTGTGTGAGCCCATCTAGGGAGCCACAGAGCGCTGGAGAGATGATGCGCATTTGG	176			
Dd	1	CAATTCTGTGTGAGCCCATCTAGGGAGCCACAGAGCGCTGGAGAGATGATGCGCATTTGG	60			
Dy	177	CCAATGGGGAATCTGTGAGAGCAAGCAAGCCCCGAGTAGAGACCCTAACCCAGCACCAA	236			
Dd	61	CCAATGGGGAATCTGTGAGAGCAAGCAAGCCCCGAGTAGAGACCCTAACCCAGCACCAA	120			
Dy	237	GAGGTAGCATTTCTGTGAGAGGACTTTCTAATAAGGGGCAATGATGCTCCCGAGGGGGTGC	296			
Dd	121	GAGGTAGCATTTCTGTGAGAGGACTTTCTAATAAGGGGCAATGATGCTCCCGAGGGGGTGC	180			
Dy	297	CTGAGCCCGCGACAGAGAGGTGCGAGGCTGGGTGCTGTCTCAATGCCCTTCAATG	356			
Dd	181	CTGAGCCCGCGACAGAGAGGTGCGAGGCTGGGTGCTGTCTCAATGCCCTTCAATG	240			

OY		357	TACCTCAACCAGCAGCTGGTGTAAACAATGCGGCCTTCGCCAGTGGAATCTGCGGACACATCTG	416
Db		241	AACCTAACCCGACGACTGTATAAATGAGGACTTTCCGACGTGGCATCTTGACACATTCTG	300
OY		417	TGAGACCGGTAACCTCCATCCTGCTCCTCTCTCTGATCATGATCTTGTGTGTGCCCC	476
Db		301	TGAGACCGGTAACCTCCATCCTGCTCCTCTCTCTGATCATGATCTTGTGTGTGCCCC	360
OY		477	TCTCTCTGCTTGAGCCTTGTCTAAGTGTGTCCCACCTGATCACGGGTGACCTGTAGGG	536
Db		361	TCCCTCCGTGTGGCTTGTCTAAGTGTGTCCCACCTGATCACGGGTGACCTGTAGGG	420
OY		537	CTGATTAGGGGTGGAGTTGTTAGAAGGAGACTTGCTGGGCTCTGTGTGTAGAGACAGCATAT	596
Db		421	CTGATTAGGGGTGGAGTTGTTAGAAGGAGACTTGCTGGGCTCTGTGTGTAGAGACAGCATAT	480
OY		597	TTGAGAGGGGATCTGTGTGTGTGCTCTGAAGGATATATCAAGAGGGGACACAGCTGTGTGT	656
Db		481	TTGAGAGGGGATCTGTGTGTGTGCTCTGAAGGATATATCAAGAGGGGACACAGCTGTGTGT	540
OY		657	TTCCCTCTTGTTGATTAACGTGAGGACAGAGGAGACGTTAGTCACGATTTTCCAANG	712
Db		541	TTCCCTCTTGTTGATTAACGTGAGGACAGAGGAGACGTTAGTCACGATTTTCCAANG	596
RESULT_15	BUTJ3254/c			
LOCUS	BUTJ3254	533 bp	mRNA	linear EST 09-OCT-2002
DEFINITION	UI-E-CII-a-fs-m-18-0-UI.s1 UI-E-CII Homo sapiens cDNA clone			
VERSION	BUTJ3254			
KEYWORDS	BUTJ3254.1 GI:23659969			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1 (bases 1 to 533) Bonaldi,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9), 791-806 (1996)			
JOURNAL MEDLINE PUBMED COMMENT	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA Tel.: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu Tissue Procurement: Dr. Gregg Hageman cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa Clone Sequencing by: Dr. M. Bento Soares, University of Iowa CDNA Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com). The following repetitive elements were found in this cDNA sequence: Sequence: 1-35, >At rich low_complexity (matched complement) Seq primer: M13 FORWARD POLIa=yes.			
FEATURES SOURCE	Location/Qualifiers  1..533 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="UI-E-CII-a-fs-m-18-0-UI" /tissue_type="RPE and Choroid" /dev_stage="adult" /lab_host="DH10B (Life Technologies) (TI phage resistant)" /clone_id="UI-E-CII" /name="Organ: eye; Vector: pVT73-Pac (Pharmacia) with a"			



ORIGIN

UT-E-C11 is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).

TAG\_TISSUE=RPE and Choroid  
TAG\_LIB=UT-E-C11  
TAG\_SEQ=ACCTA"

Query Match 36.2%; Score 431; DB 13; Length 533;  
Best Local Similarity 99.6%; Pred. No. 1e-77;  
Matches 531; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	544	GAGTGGGTTTGTGAAGGACCTGCTGGGCTTGGTGTGAGACGACATATTGGAGG	603
DB	533	GAGTGGGTTTGTGAAGGACCTGCTGGGCTTGGTGTGAGACGACATATTGGAGG	474
QY	604	GATCTGGTGGTGGCTTGAAGGATATGAGAGGGGACACAGTGTGTTCCTTCCCT	663
DB	473	GATCTGGTGGTGGCTTGAAGGATATGAGAGGGGACACAGTGTGTTCCTTCCCT	414
QY	664	TTGTGTTAACCGTGAAGGACGATGATGTCAGCATTTCCAAAGTGGGTGGGT	723
DB	413	TTGTGTTAACCGTGAAGGACGATGATGTCAGCATTTCCAAAGTGGGTGGGT	354
QY	724	CCGTGGTCCCGAGTACTTTTAAAGTGTATGGGCTGCACTAAAGTGGCAAAATCA	783
DB	353	CCGTGGTCCCGAGTACTTTTAAAGTGTATGGGCTGCACTAAAGTGGCAAAATCA	294
QY	784	GAGCAGAAAGCATGCTCCCAATTCCTCAATCCTTTATGCCGAGAGATCTCAG	843
DB	293	GAGCAGAAAGCATGCTCCCAATTCCTCAATCCTTTATGCCGAGAGATCTCAG	234
QY	844	CTGGATGCCAAGTGTCCGATGCTGTGAAAGACATGCCAGTCTCCTGCTAGGG	903
DB	233	CTGGATGCCAAGTGTCCGATGCTGTGAAAGACATGCCAGTCTCCTGCTAGGG	174
QY	904	AGCAGGACTTGGGCTTAGGGGAGGTGAGAAAATTCACACTTTTATGCACTGTTTG	963
DB	173	AGCAGGACTTGGGCTTAGGGGAGGTGAGAAAATTCACACTTTTATGCACTGTTTG	114
QY	964	TTTAAATGATATTTTATTTGCTACTTATTTAGTGTAGACAGTGTAGTGCATTCT	1023
DB	113	TTTAAATGATATTTTATTTGCTACTTATTTAGTGTAGACAGTGTAGTGCATTCT	54
QY	1024	ATTATTTGACCTTTTCAATATAGATTAAATTAATTAATTAATTAATTAATTAATTA	1076
DB	53	ATTATTTGACCTTTTCAATATAGATTAAATTAATTAATTAATTAATTAATTAATTA	1

Search completed: August 22, 2004, 07:09:17  
Job time : 3000 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: August 22, 2004, 04:42:00 ; Search time 117 Seconds

(without alignments)  
5653.859 Million cell updates/sec

Title: US-10-001-885-31

Sequence: 1 aaacgccttgagccgcgcg.....gaagagcacacggaacc 1192

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/prodata/2/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/prodata/2/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/prodata/2/ina/6A.COMB.seq:\*
- 4: /cgn2\_6/prodata/2/ina/6B.COMB.seq:\*
- 5: /cgn2\_6/prodata/2/ina/PCTUS.COMB.seq:\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	792	66.4	US-09-149-476-58	Sequence 58, Appl
2	473	39.7	US-09-621-976-864	Sequence 864, App
3	185	15.5	US-09-621-976-3064	Sequence 3064, Ap
4	34	2.9	US-08-747-221B-24	Sequence 24, Appl
5	34	2.9	US-08-747-221B-26	Sequence 26, Appl
6	34	2.9	US-09-005-051-24	Sequence 24, Appl
7	34	2.9	US-09-005-051-26	Sequence 26, Appl
8	34	2.9	US-09-403-942F-24	Sequence 24, Appl
9	34	2.9	US-09-403-942F-26	Sequence 26, Appl
10	32	2.7	US-09-621-976-17202	Sequence 17202, A
11	32	2.7	US-09-621-976-17202	Sequence 17202, A
12	32	2.7	US-08-368-776A-4	Sequence 4, Appl
13	32	2.7	PCT-US96-00419-4	Sequence 4, Appl
14	32	2.7	US-09-918-686-2	Sequence 2, Appl
15	32	2.7	US-09-918-686-1	Sequence 1, Appl
16	31	2.6	US-09-328-111-797	Sequence 797, App
17	31	2.6	US-09-328-111-125	Sequence 125, App
18	31	2.6	US-09-713-550-125	Sequence 125, App
19	31	2.6	US-09-328-111-436	Sequence 436, App
20	31	2.6	US-09-328-111-561	Sequence 561, App
21	31	2.6	US-08-757-046A-5	Sequence 5, Appl
22	31	2.6	US-09-447-208A-5	Sequence 5, Appl
23	31	2.6	US-09-135-988-5	Sequence 5, Appl
24	31	2.6	US-09-277-716-5	Sequence 5, Appl
25	31	2.6	US-08-587-274A-5	Sequence 5, Appl
26	31	2.6	US-08-908-909-5	Sequence 5, Appl
27	31	2.6	US-09-609-161B-5	Sequence 5, Appl

28	31	2.6	US-08-990-103-5	Sequence 5, Appl
29	31	2.6	US-09-746-485A-5	Sequence 5, Appl
30	31	2.6	US-10-126-139-5	Sequence 5, Appl
31	31	2.6	US-10-126-798-5	Sequence 5, Appl
32	31	2.6	US-10-126-777-5	Sequence 5, Appl
33	31	2.6	US-09-426-557-1	Sequence 1, Appl
34	31	2.6	US-08-488-199-3	Sequence 3, Appl
35	31	2.6	US-08-147-777-3	Sequence 3, Appl
36	31	2.6	US-08-452-872-3	Sequence 3, Appl
37	31	2.6	PCT-US93-03985-3	Sequence 3, Appl
38	31	2.6	US-09-497-855A-37	Sequence 37, Appl
39	31	2.6	US-09-497-855A-44	Sequence 44, Appl
40	30	2.5	US-09-197-814-9	Sequence 9, Appl
41	30	2.5	US-09-920-581-9	Sequence 9, Appl
42	30	2.5	US-09-165-239A-5	Sequence 5, Appl
43	30	2.5	US-09-894-916-3	Sequence 3, Appl
44	30	2.5	US-09-404-879A-201	Sequence 201, App
45	30	2.5	US-09-338-933-201	Sequence 201, App

## ALIGNMENTS

RESULT 1  
US-09-149-476-58  
Sequence 58, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION: a1.  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
EARLIER FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23

**Tue Aug 24 09:48:32 2004**

us-10-001-885-31.oli.rni

Page 2

[illegible][illegible]

EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 66.4%; Score 792; DB 4; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0; Mismatches 0; Indels 2; Gaps 2;  
Matches 1032; Conservative 0

46 TCGCCCTGACAGACAGATCTACTACGCGTGTCACTTGTGAACATCACTGACTG 105  
1 TCGCCCTGACAGACAGATCTACTACGCGTGTCACTTGTGAACATCACTGACTG 60  
106 CAAGCTCTCTCAATTTCTGTGTGAGCCCATCAGAGGACCCACAGGCGCTGGAGAGATGT 165  
61 CAAGCTCTCTCAATTTCTGTGTGAGCCCATCAGAGGACCCACAGGCGCTGGAGAGATGT 120  
166 GCGGATCTTGCCCAATGAGGGAATGTGTGAGGACGACGACCCCGAGTGAAGACCACTAC 225  
121 GCGGATCTTGCCCAATGAGGGAATGTGTGAGGACGACGACCCCGAGTGAAGACCACTAC 180  
226 CCAAGCCACCAAGAGTGTGATCTCTGACAGAGCTTTCTGAATAGGGGCAATGTGTCTCC 285  
181 CCAAGCCACCAAGAGTGTGATCTCTGACAGAGCTTTCTGAATAGGGGCAATGTGTCTCC 240  
286 CCAAGGCGGCTCTGAGCCCGGACAGACAGGACAGGCTGAGTGTGTCTGCTCACTC 345  
241 CCAAGGCGGCTCTGAGCCCGGACAGACAGGACAGGCTGAGTGTGTCTGCTCACTC 300  
346 CCGCTTCAATGACCTCAACCGGACGCTGTGAACATGAGCTTTCCGACAGTGTGACTG 405  
301 CCGCTTCAATGACCTCAACCGGACGCTGTGAACATGAGCTTTCCGACAGTGTGACTG 360  
406 CAACCATGCTGTGAGACCGGTGTACCTCCATCCGTGCTCTCTCTGCTCATGATGCTTGG 465  
361 CAACCATGCTGTGAGACCGGTGTACCTCCATCCGTGCTCTCTCTGCTCATGATGCTTGG 420  
466 TGTTCGTGCGCTCTCTGCTGTGAGCTTTGTAAGCTGTGTGCTCACTGATGCTGAGT 525  
421 TGTTCGTGCGCTCTCTGCTGTGAGCTTTGTAAGCTGTGTGCTCACTGATGCTGAGT 480  
526 ACCTGTGAGGCGGTGTGAGGCGGTGTGTAAGAGGACCTTGTGAGGCGGTGTGAGT 585  
481 ACCTGTGAGGCGGTGTGAGGCGGTGTGTAAGAGGACCTTGTGAGGCGGTGTGAGT 540  
586 AGAGAGCATTTTGAAGGAGGATCTGTGTGAGCTTGAAGGATGATCAGAGAGGAGCA 645  
541 AGAGAGCATTTTGAAGGAGGATCTGTGTGAGCTTGAAGGATGATCAGAGAGGAGCA 600  
646 CAGGTGTGTCTTCCCTTTGTGTGAAGCTGTGAGGACAGGACCTTGAAT - CCAAGCAT 704  
601 CAGGTGTGTCTTCCCTTTGTGTGAAGCTGTGAGGACAGGACCTTGAAT - CCAAGCAT 660  
705 TT - CCAAGCTGTGAGGCGGTGTGAGGCGGTGTGTAAGGATGATGAGGCGGTG 763  
661 TTCCCAAGAGTGTGAGGCGGTGTGAGGCGGTGTGTAAGGATGATGAGGCGGTG 720  
764 CATTAAGTGCACAAATCAGAGCAAGAAAGCATGCTTCCCAATCTCTCAATCTCT 823  
721 CATTAAGTGCACAAATCAGAGCAAGAAAGCATGCTTCCCAATCTCTCAATCTCT 780  
824 TTATGCGGAGAGATCTCAGCTGTGATGCAATGTTCCGATGCTGTGAGAAAGATGCC 883  
781 TTATGCGGAGAGATCTCAGCTGTGATGCAATGTTCCGATGCTGTGAGAAAGATGCC 840  
884 GAGCTCTCTGCTGCTGAGGAGAGGACCTTGTGAGGCGGTGTGAGGCGGTGTGAG 943  
841 GAGCTCTCTGCTGCTGAGGAGAGGACCTTGTGAGGCGGTGTGAGGCGGTGTGAG 900  
944 CTTTGTAGCATGCTTTTGTGTTTAAATGATATTTTATGCTACTTTATGTTAGG 1003

901 CTTTGTAGCATGCTTTTGTGTTTAAATGATATTTTATGCTACTTTATGTTAGG 960  
1004 ACAAGTGTGAGTGTGATCTATTTATTTATGACCTTTTCAATAAATGATTAAGTAAAA 1063  
961 ACAAGTGTGAGTGTGATCTATTTATTTATGACCTTTTCAATAAATGATTAAGTAAAA 1020  
1064 AAAAAAAAAAAAAA 1077  
1021 AAAAAAAAAAAAAA 1034

RESULT 2  
US-09-621-976-864  
Sequence 864, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J. B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J. Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET-054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent pm  
SEQ ID NO 864  
LENGTH: 483  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURES:  
NAME/KEY: CDS  
LOCATION: 171..482  
US-09-621-976-864

Query Match 39.7%; Score 473; DB 4; Length 483;  
Best Local Similarity 100.0%; Pred. No. 9,6e-202; Indels 0; Gaps 0;  
Matches 473; Conservative 0; Mismatches 0

1 AAACCGCTGAGCGCGCGGAGTGAAGCGCGGAGGCGCGGAGTGCCTGACAGCA 60  
11 AAACCGCTGAGCGCGCGGAGTGAAGCGCGGAGGCGCGGAGTGCCTGACAGCA 70  
61 CAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACTGCAAGCTCTCTCAAT 120  
71 CAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACTGCAAGCTCTCTCAAT 130  
121 TTCTGTGACAGCCCATCAGGAGACCCACAGCGCTGTGAGGAGATGTGAGGATCTTGGCCAA 180  
131 TTCTGTGACAGCCCATCAGGAGACCCACAGCGCTGTGAGGAGATGTGAGGATCTTGGCCAA 190  
181 TGGGGAATGCTGACAGAGCAAGACCCCGAGTGAAGCACTAACCCAGCAAGAGG 240  
191 TGGGGAATGCTGACAGAGCAAGACCCCGAGTGAAGCACTAACCCAGCAAGAGG 250  
241 TAGCATCTCTGACAGAGCTTTCTTAATAGGAGCCATGATGCTCTCCCAAGGCGGTCTGG 300  
251 TAGCATCTCTGACAGAGCTTTCTTAATAGGAGCCATGATGCTCTCCCAAGGCGGTCTGG 310  
301 CCGCGCGAGAGAGAGGAGGAGGCTGAGGCTGCTGCTGAGTCCCTTCAATGACT 360  
311 CCGCGCGAGAGAGAGGAGGAGGCTGAGGCTGCTGCTGAGTCCCTTCAATGACT 370  
361 CAACCGGAGCTGTGAACATGAGGCTTTCCGAGTGTGATCTCCGCAACATGCTGTGA 420  
371 CAACCGGAGCTGTGAACATGAGGCTTTCCGAGTGTGATCTCCGCAACATGCTGTGA 430  
421 GCGGAGTCACTCACTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 473  
431 GCGGAGTCACTCACTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483

RESULT 3  
US-09-621-976-3064

Sequence 3064, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET 054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.Pm  
SEQ ID NO 3064  
LENGTH: 519  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 93..431  
US-09-621-976-3064

Query Match 15.5%; Score 185; DB 4; Length 519;  
Best Local Similarity 99.6%; Pred. No. 3.2e-73;  
Matches 235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 58 ACAAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACGAGGCTCCCTC 117  
DB 116 ACACAGCATCTACTCAGCGTGGGTCACTCTGTGAACATCACTGACGAGGCTCCCTC 175  
QY 118 AATTTCGTGTCAGCCCATAGGACCCACAGGCGCTGGAGAGTGTGCGATCTTGGC 177  
DB 176 AATTTCGTGTCAGCCCATAGGACCCACAGGCGCTGGAGAGTGTGCGATCTTGGC 235  
QY 178 CAATGGGGAATGTCGTGAGGACGACGACCCCGAGTGAAGACCACTACCCACCCACG 237  
DB 236 CAATGGGGAATGTCGTGAGGACGACGACCCCGAGTGAAGACCACTACCCACCCACG 295  
QY 238 AGGTAGCATCTCTGCAAGAGCTTCTTCAATAGGGGCAATGTGTCTCCCGCAGGG 293  
DB 296 AGGTAGCATCTCTGCAAGAGCTTCTTCAATAGGGGCAATGTGTCTCCCGCAGGG 351

## RESULT 4

US-08-747-221B-24  
Sequence 24, Application US/08747221B  
Patent No. 6063610

GENERAL INFORMATION:  
APPLICANT: Silver, Gary W.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: No. 6063610e1 Carboxylesterase Nucleic Acid  
TITLE OF INVENTION: Molecules, Proteins and Uses Thereof  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/747,221B  
FILING DATE: No. 6063610e1 December 12, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: FC-1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2836 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 99..1889  
US-08-747-221B-24

Query Match 2.9%; Score 34; DB 3; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 7.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TTTAAGTAAAAA  
DB 2778 TTTAAGTAAAAA  
1085  
2811

## RESULT 5

US-08-747-221B-26/C  
Sequence 26, Application US/08747221B  
Patent No. 6063610

GENERAL INFORMATION:  
APPLICANT: Silver, Gary W.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: No. 6063610e1 Carboxylesterase Nucleic Acid  
TITLE OF INVENTION: Molecules, Proteins and Uses Thereof  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/747,221B  
FILING DATE: No. 6063610e1 December 12, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: FC-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2836 nucleotides  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-747-221B-26

Query Match 2.9%; Score 34; DB 3; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 7.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TTTAAGTAAAAA  
1085  
2811

Db 59 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 26

RESULT 6  
US-09-005-051-24  
Sequence 24, Application US/09005051

PATENT INFORMATION:

APPLICANT: Silver, Gary W.

APPLICANT: Wisniewski, Nancy

TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid

NUMBER OF SEQUENCES: 66

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carol Talkington Verser, Ph.D.

ADDRESSEE: Heska Corporation

STREET: 1825 Sharp Point Drive

CITY: Fort Collins

STATE: Colorado

COUNTRY: USA

ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: Wordperfect for Windows, Version 7.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/005,051

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/747,221

FILING DATE: NO. 6291222el December 12, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Verser, Carol Talkington

REGISTRATION NUMBER: 37,459

REFERENCE/DOCKET NUMBER: FC-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 970/493-7272

TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 2836 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 99...1889

US-09-005-051-24

Query Match 2.9%; Score 34; DB 3; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 7.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1052 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 1085  
2778 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 2811

RESULT 7  
US-09-005-051-26/c  
Sequence 26, Application US/09005051

PATENT INFORMATION:

GENERAL INFORMATION:

APPLICANT: Silver, Gary W.

APPLICANT: Wisniewski, Nancy

TITLE OF INVENTION: No. 6291222el Carboxylesterase Nucleic Acid

NUMBER OF SEQUENCES: 66

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carol Talkington Verser, Ph.D.

ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: Wordperfect for Windows, Version 7.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/005,051

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/747,221

FILING DATE: NO. 6291222el December 12, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Verser, Carol Talkington

REGISTRATION NUMBER: 37,459

REFERENCE/DOCKET NUMBER: FC-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 970/493-7272

TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 2836 nucleotides

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-005-051-26

Query Match 2.9%; Score 34; DB 3; Length 2836;  
Best Local Similarity 100.0%; Pred. No. 7.4e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1052 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 1085  
59 TTTAAGTAAAAAAAAAAAAAAAAAAAAA 26

RESULT 8  
US-09-403-942F-24  
Sequence 24, Application US/09403942F

PATENT INFORMATION:

GENERAL INFORMATION:

APPLICANT: Silver, Gary W.

APPLICANT: Brandt, Kevin S.

TITLE OF INVENTION: NOVEL CARBOXYLESTERASE NUCLEIC ACID MOLECULES, PROTEINS AND USES

FILE REFERENCE: FC-1-CI-PUS

CURRENT APPLICATION NUMBER: US/09/403,942F

PRIOR FILING DATE: 2000-05-02

PRIOR APPLICATION NUMBER: PCT/US97/20598

PRIOR FILING DATE: 1997-11-10

PRIOR APPLICATION NUMBER: 08/747,221

PRIOR FILING DATE: 1996-11-12

NUMBER OF SEQ ID NOS: 76

SOFTWARE: Patent version 3.1

SEQ ID NO 24

LENGTH: 2836

TYPE: DNA

ORGANISM: Ctenocephalides felis

FEATURE:

NAME/KEY: CDS

LOCATION: (99)..(1889)

OTHER INFORMATION:

FEATURE: misc\_feature

NAME/KEY: misc\_feature

LOCATION: (2278)..(2278)

OTHER INFORMATION: n = unknown





REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2901 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic  
US-08-368-776A-4

Query Match 2.7%; Score 32; DB 4; Length 2901;  
Best Local Similarity 100.0%; Pred. No. 5.8e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 TAAGTAAAAAAAAAAAAAAAAAAAAAAAAA 1085  
DB 2853 TAAGTAAAAAAAAAAAAAAAAAAAAAAAAA 2884

RESULT 13  
PCT-US96-00419-4  
Sequence 4, Application PC/TUS9600419  
GENERAL INFORMATION:  
APPLICANT: Thomas Closek, Axel Ullrich, Birgit  
TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF  
TITLE OF INVENTION: MDX1 SIGNAL TRANSDUCTION DISORDERS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/00419  
FILING DATE: January 3, 1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: none  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 208/007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2901  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic  
PCT-US96-00419-4

Query Match 2.7%; Score 32; DB 5; Length 2901;  
Best Local Similarity 100.0%; Pred. No. 5.8e-05;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1054 TAAGTAAAAAAAAAAAAAAAAAAAAAAAAA 1085  
DB 2853 TAAGTAAAAAAAAAAAAAAAAAAAAAAAAA 2884

RESULT 14  
US-09-918-686-2  
Sequence 2, Application US/09918666  
Patent No. 6475739  
GENERAL INFORMATION:  
APPLICANT: Brunkow, Mary  
APPLICANT: Prohl, Sean  
APPLICANT: Paepfer, Bryan  
APPLICANT: Staehling-Hampton, Karen  
TITLE OF INVENTION: METHODS FOR IDENTIFYING  
TITLE OF INVENTION: GENOMIC DELETIONS  
FILE REFERENCE: 240083.515  
CURRENT APPLICATION NUMBER: US/09/918,686  
CURRENT FILING DATE: 2001-07-30  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 51719  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1246..2572, 2604  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-686-2

Query Match 2.7%; Score 32; DB 4; Length 51719;  
Best Local Similarity 100.0%; Pred. No. 5.3e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAG 1086  
DB 44873 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAG 44904

RESULT 15  
US-09-918-686-1  
Sequence 1, Application US/09918666  
Patent No. 6475739  
GENERAL INFORMATION:  
APPLICANT: Brunkow, Mary  
APPLICANT: Prohl, Sean  
APPLICANT: Paepfer, Bryan  
APPLICANT: Staehling-Hampton, Karen  
TITLE OF INVENTION: METHODS FOR IDENTIFYING  
TITLE OF INVENTION: GENOMIC DELETIONS  
FILE REFERENCE: 240083.515  
CURRENT APPLICATION NUMBER: US/09/918,686  
CURRENT FILING DATE: 2001-07-30  
NUMBER OF SEQ ID NOS: 105  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 92139  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 7043..8369, 8401  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-686-1

Query Match 2.7%; Score 32; DB 4; Length 92139;  
Best Local Similarity 100.0%; Pred. No. 5.2e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1055 AAGTAAAAAAAAAAAAAAAAAAAAAAAAAG 1086

Tue Aug 24 09:48:32 2004

us-10-001-885-31.011.rn1

Page 8

Db 50660 AACTTAAAAAAAAAAAAAAAAAG 50691

Search completed: August 22, 2004, 07:11:24  
Job time : 119 secs

OM nucleic - nucleic search, using sw model

Run on: August 22, 2004, 06:19:16 ; Search time 594 Seconds

(without alignments)  
9857.345 Million cell updates/sec

Title: US-10-001-885-31

Sequence: 1 aaacgccttgagagccgcgcg.....gaagagagacacaggaaccc 1192

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 3228839 seqs, 245606551 residues

Word size: 0

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCRUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1192	100.0	1192	13	US-10-001-885-31
2	903	75.8	1108	16	US-10-264-237-619
3	792	66.4	1049	10	US-09-809-331-58
4	792	66.4	1049	10	US-09-882-171-58
5	792	66.4	1049	13	US-10-164-861-58
6	669	56.1	1052	13	US-10-302-172-94
7	326	27.3	365	15	US-10-076-747-65
8	195	16.4	612	15	US-10-076-747-64
9	125	10.5	643	13	US-10-001-885-30
10	38	3.2	473	9	US-09-864-761-551
11	38	3.2	473	9	US-09-864-761-521
12	38	3.2	998	13	US-10-027-632-209023
13	38	3.2	998	16	US-10-027-632-209023
14	34	2.9	616	13	US-10-424-599-125925

15	34	2.9	2836	17	US-10-678-521-24	Sequence 24, Appl
16	34	2.9	2836	17	US-10-678-521-26	Sequence 26, Appl
17	33	2.8	653	12	US-09-969-034-718	Sequence 718, Appl
18	33	2.8	758	13	US-10-027-632-127811	Sequence 127811,
19	33	2.8	758	16	US-10-027-632-127811	Sequence 127811,
20	33	2.8	833	15	US-10-198-846-5140	Sequence 5140, Ap
21	33	2.8	1473	17	US-10-437-863-99876	Sequence 99876, A
22	33	2.8	1885	16	US-10-374-780A-849	Sequence 849, App
23	33	2.8	96595	12	US-09-997-722-202	Sequence 202, App
24	33	2.8	21231	13	US-10-087-192-1126	Sequence 1126, Ap
25	33	2.8	3673778	15	US-10-312-841-2	Sequence 2, Appl
26	32	2.7	152	9	US-09-834-975-110	Sequence 110, App
27	32	2.7	157	9	US-09-867-701-9937	Sequence 9937, Ap
28	32	2.7	389	17	US-10-437-863-26389	Sequence 26389, A
29	32	2.7	473	16	US-10-060-036-916	Sequence 916, App
30	32	2.7	480	15	US-10-131-827-8192	Sequence 8192, App
31	32	2.7	572	17	US-10-437-863-88429	Sequence 88429, A
32	32	2.7	602	9	US-09-918-686-7	Sequence 7, Appl
33	32	2.7	602	15	US-10-353-150-7	Sequence 7, Appl
34	32	2.7	649	15	US-10-066-543-727	Sequence 727, App
35	32	2.7	706	15	US-10-050-704-62	Sequence 62, Appl
36	32	2.7	706	17	US-10-798-512-62	Sequence 62, Appl
37	32	2.7	1525	9	US-09-880-107-3816	Sequence 3816, Appl
38	32	2.7	1852	17	US-10-437-863-5371	Sequence 5371, Ap
39	32	2.7	2138	10	US-09-814-353-19280	Sequence 19280, A
40	32	2.7	3031	13	US-10-027-632-113397	Sequence 113397,
41	32	2.7	3031	16	US-10-027-632-113397	Sequence 113397,
42	32	2.7	3685	9	US-09-822-830A-331	Sequence 331, App
43	32	2.7	5204	15	US-10-311-455-873	Sequence 873, App
44	32	2.7	33112	16	US-10-429-873A-3	Sequence 3, Appl
45	32	2.7	39001	16	US-10-199-199-14	Sequence 14, Appl

## ALIGNMENTS

```
RESULT 1
US-10-001-885-31
; Sequence 31, Application US/10001885
; Publication No US20040058319A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Caffarely, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; FILE REFERENCE: DEX-0279
; CURRENT APPLICATION NUMBER: US/10/001,885
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,061
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/253,257
; NUMBER OF SEQ ID NOS: 167
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 1192
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-001-885-31
Query Match 100.0%; Score 1192; DB 13; Length 1192;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AAACCGCTGAGAGCCGCGGAGAGACCGCCGAGAGCCGCGGAGTGGCGGCTGACAGACA 60
1 AAACCGCTGAGAGCCGCGGAGAGACCGCCGAGAGCCGCGGAGTGGCGGCTGACAGACA 60
CAGCATCTACCTCAGCGTGGTGCACCTCTGTGAACATGACTGACAGAGCTCCCTCAAT 120
```

Db 61 CAGCATCTACAGCGGAGTCACTCTGTGAACATCACTGACTGCAAGCTCCCTCAAT 120  
QY 121 TTCTGTGCGAGCCCATCAGAGGACCCCAAGGCGCTGGAGAGATGCTGGCCCA 180  
Db 121 TTCTGTGCGAGCCCATCAGAGGACCCCAAGGCGCTGGAGAGATGCTGGCCCA 180  
QY 181 TGGGGAATATGTGCGAGAGACAGACCCCGAGTGAAGACCACTACCCAGCCCAAGAG 240  
Db 181 TGGGGAATATGTGCGAGAGACAGACCCCGAGTGAAGACCACTACCCAGCCCAAGAG 240  
QY 241 TAGCATCTCTGACAGAGCTTCTTCAATAGGGCCATGCTGCTCCCGAGGGGCTCTGG 300  
Db 241 TAGCATCTCTGACAGAGCTTCTTCAATAGGGCCATGCTGCTCCCGAGGGGCTCTGG 300  
QY 301 CCCCCCGCAGACAGAGAGAGTGCAGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
Db 301 CCCCCCGCAGACAGAGAGAGTGCAGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
QY 361 CAACCCGAGCTGTGAACATGAGGCTTTCCGAGTGGCAATCTCGCAACCAATGCTGTGA 420  
Db 361 CAACCCGAGCTGTGAACATGAGGCTTTCCGAGTGGCAATCTCGCAACCAATGCTGTGA 420  
QY 421 GCGGTGACCTCCATCTGCT 480  
Db 421 GCGGTGACCTCCATCTGCT 480  
QY 481 CCTGTGTGCTGTCT 540  
Db 481 CCTGTGTGCTGTCT 540  
QY 541 TAGGGGCTGCTGTGTGAAGGAGCTTGTGAGGAGCTTGTGAGGAGGAGGAGGAGGAG 600  
Db 541 TAGGGGCTGCTGTGTGAAGGAGCTTGTGAGGAGCTTGTGAGGAGGAGGAGGAGGAG 600  
QY 601 AGGGGATCTGTGTGTGAAGGAGCTTGTGAAGTATGATCAGAGAGGAGGAGGAGGAG 660  
Db 601 AGGGGATCTGTGTGTGAAGGAGCTTGTGAAGTATGATCAGAGAGGAGGAGGAGGAG 660  
QY 661 CCTTTGTGTAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
Db 661 CCTTTGTGTAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
QY 721 GGTGCGTGTGCTCCGAGTACTTCTTGAAGTATGAGTATGAGGAGGAGGAGGAGGAG 780  
Db 721 GGTGCGTGTGCTCCGAGTACTTCTTGAAGTATGAGTATGAGGAGGAGGAGGAGGAG 780  
QY 781 TCAAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
Db 781 TCAAGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
QY 841 CAGCTGATATGCAACATGCTTCCGATGCTGTGAGACATGCTGCTGCTGCTGCTGCT 900  
Db 841 CAGCTGATATGCAACATGCTTCCGATGCTGTGAGACATGCTGCTGCTGCTGCTGCT 900  
QY 901 GAGAGCAGAGCTTGGGCTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
Db 901 GAGAGCAGAGCTTGGGCTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
QY 961 TTGTTTAAATGATATTTTAAATGATATTTTAAATGATATTTTAAATGATATTTTAA 1020  
Db 961 TTGTTTAAATGATATTTTAAATGATATTTTAAATGATATTTTAAATGATATTTTAA 1020  
QY 1021 TCAATTAATGATATTTTAAATGATATTTTAAATGATATTTTAAATGATATTTTAA 1080  
Db 1021 TCAATTAATGATATTTTAAATGATATTTTAAATGATATTTTAAATGATATTTTAA 1080  
QY 1081 AAAAAGCTGTGGGAGTTACCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
Db 1081 AAAAAGCTGTGGGAGTTACCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140  
QY 1141 GCGCCCAATTTCCCGCAATTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1192  
Db 1141 GCGCCCAATTTCCCGCAATTTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1192

RESULT 2  
US-10-264-237-619  
; Sequence 619, Application US/10264237  
; Publication No. US20040039491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birex et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P131p1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; PRIOR FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 619  
; LENGTH: 1108  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURES:  
; NAME/KEY: misc feature  
; LOCATION: (1101)..(1101)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-264-237-619  
Query Match 75.8%; Score 903; DB 16; Length 1108;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1053; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 22 AGTGAAGCGCCCGCAGAGCCCGGAGATGCGGCTGTGAGACAGACATCTACTACGCTGGGT 81  
Db 8 AGTGAAGCGCCCGCAGAGCCCGGAGATGCGGCTGTGAGACAGACATCTACTACGCTGGGT 67  
QY 82 CACCTGTGTAACATCACTGACTGCAAGCTCCCTCAATTTCTGTGTCAGACCCATCAGG 141  
Db 68 CACCTGTGTAACATCACTGACTGCAAGCTCCCTCAATTTCTGTGTCAGACCCATCAGG 127  
QY 142 ACCCAAGCGCTGTGAGAGATGTGCGGATCTTGGCCATGCGGAAATGCTGAGAGCA 201  
Db 128 ACCCAAGCGCTGTGAGAGATGTGCGGATCTTGGCCATGCGGAAATGCTGAGAGCA 187  
QY 202 CGACCCCGAGTGAAGGAGCACTACCCAGCCAGCAAGAGTAGCATTCCTGAGAGACTT 261  
Db 188 CGACCCCGAGTGAAGGAGCACTACCCAGCCAGCAAGAGTAGCATTCCTGAGAGACTT 247  
QY 262 CTTCATATAGGGCCATGATGTCTCCCGAGGGGATCTGGCCCCCGCAGAGAGGAGG 321  
Db 248 CTTCATATAGGGCCATGATGTCTCCCGAGGGGATCTGGCCCCCGCAGAGAGGAGG 307  
QY 322 TCCAGAGCTGGGTCTGTCTAGTCCCTTCAATGACCTTAACCGAGAGCTGTGAAAT 381  
Db 308 TCCAGAGCTGGGTCTGTCTAGTCCCTTCAATGACCTTAACCGAGAGCTGTGAAAT 367  
QY 382 GGGCTTTCGAGAGGAGCATCGGCAACATGCTGTGAGAGCCGAGTGAATCTGCT 441  
Db 368 GGGCTTTCGAGAGGAGCATCGGCAACATGCTGTGAGAGCCGAGTGAATCTGCT 427  
QY 442 CCTCTTCTGTCTATGATGCTGTGTGTGCTGCTCTCTCTGCTGTGCTGTGCT 501  
Db 428 CCTCTTCTGTCTATGATGCTGTGTGTGCTGCTCTCTCTGCTGTGCTGTGCT 487  
QY 502 GATGTCACCTGATGAGCGGTGACCTGAGAGGCTGATAGGGGTGCTTTGTGAGAG 561  
Db 488 GATGTCACCTGATGAGCGGTGACCTGAGAGGCTGATAGGGGTGCTTTGTGAGAG 547  
QY 562 GACTGTGTGGGCTGTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 621  
Db 548 GACTGTGTGGGCTGTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 607  
QY 622 AAGATGATCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 681

Db 608 AAGGTATGATCGAGAGAGGAGACACAGAGGTGTGTTTCCCTTTGTGTTAAAGCGTAGAGC 667  
Qy 682 AGAGGAGACGTATGTCAGACATTTCCAAAGTGAGGTGGTCCGTTGGTTCCGAGATA 741  
Db 668 AAGGAGAGACGTTATGTCAGACATTTCCAAAGTGAGGTGGTCCGTTGGTTCCGAGATA 727  
Qy 742 CTTTATGAGTGAGGTGGTCCGTTATGAGTGCACAAATCAGACAGAGAGAGAGATGCC 801  
Db 728 CTTTATGAGTGAGGTGGTCCGTTATGAGTGCACAAATCAGACAGAGAGAGATGCC 787  
Qy 802 CTTCCCAATTTCTCAATCTTTTATGCGAGAGATCTCAGCTGGAGTCCCAATGTTTC 861  
Db 788 CTTCCCAATTTCTCAATCTTTTATGCGAGAGATCTCAGCTGGAGTCCCAATGTTTC 847  
Qy 862 CGATGCTGTGAGAGACATGCCAGCTCTCTCTGCTTATGAGAGAGAGACTTGGCTTAG 921  
Db 848 CGATGCTGTGAGAGACATGCCAGCTCTCTCTGCTTATGAGAGAGACTTGGCTTAG 907  
Qy 922 GCGAGGTGAGAGAGATTTCCAGACTTTTATGAGAGAGAGAGAGAGAGAGAGAGAG 981  
Db 908 GCGAGGTGAGAGAGATTTCCAGACTTTTATGAGAGAGAGAGAGAGAGAGAGAGAG 967  
Qy 982 ATTGGCTACTTTATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1041  
Db 968 ATTGGCTACTTTATGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1027  
Qy 1042 AATAAATAGATTAAAGTAAAAAATAAAAAATAAAAA 1077  
Db 1028 AATAAATAGATTAAAGTAAAAAATAAAAAATAAAAA 1063

RESULT 3  
US-09-809-391-58  
; Sequence 58, Application US/09809391  
; Publication No. US20030049618A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/809,391  
; CURRENT FILING DATE: 2001-03-16  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 761  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 58  
; LENGTH: 1049  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-809-391-58

Query Match 66.4%; Score 792; DB 10; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
Qy 46 TGGCGCTGTCAGACAGATCTACTCAGCGTGGTCACTCTGTGAACATCATGACTG 105  
Db 1 TGGCGCTGTCAGACAGATCTACTCAGCGTGGTCACTCTGTGAACATCATGACTG 60  
Qy 106 CAAGCTCTCCCTCAATTTCTGTGACGCCATCAGGAGCCCAAGCGCTGGAGAGATGT 165  
Db 61 CAAGCTCTCCCTCAATTTCTGTGACGCCATCAGGAGCCCAAGCGCTGGAGAGATGT 120  
Qy 166 GCGAGTCTTGCCCAATGCGGAAATCGTGCAGACACACCCCGAGTGAAGACACTAC 225  
Db 121 GCGAGTCTTGCCCAATGCGGAAATCGTGCAGACACACCCCGAGTGAAGACACTAC 180  
Qy 226 CCAGCCACCAAGAGAGATCTCTGACAGAGCTTCTTCAATAGAGGGCATGATGCTCC 285  
Db 181 CCAGCCACCAAGAGAGATCTCTGACAGAGCTTCTTCAATAGAGGGCATGATGCTCC 240  
Qy 286 CCCAGGGGAGTCTGAGCCCGCCAGACAGAGGAGGCTGAGGCTGCTCACTC 345

Db 241 CCCAGGGGAGTCTGAGCCCGCCAGACAGAGAGTGCAGAGCTGGGTCTGCTACGT 300  
Qy 346 CCCCTTCAATGACCTCAACCGGAGCTGTGTGAACATGAGGCTTTCCGAGTGGCATCTCG 405  
Db 301 CCCCTTCAATGACCTCAACCGGAGCTGTGTGAACATGAGGCTTTCCGAGTGGCATCTCG 360  
Qy 406 CAACATGCTGTGAGAGCCCGGTGACCTCATCTGCTCTCTCTCTCATGATGCTTGG 465  
Db 361 CAACATGCTGTGAGAGCCCGGTGACCTCATCTGCTCTCTCTCTCATGATGCTTGG 420  
Qy 466 TGTGTGAGGCTCTCTGAGTGGCTGTGTCTACCTGTGTGCCACCTGTGAGGCTG 525  
Db 421 TGTGTGAGGCTCTCTGAGTGGCTGTGTCTACCTGTGTGCCACCTGTGAGGCTG 480  
Qy 526 ACCTTGAAGGCTGTATAGGAGTGGGTTTGTGAAGGAGACTTGTGGCTTGGTGTAG 585  
Db 481 ACCTTGAAGGCTGTATAGGAGTGGGTTTGTGAAGGAGACTTGTGGCTTGGTGTAG 540  
Qy 586 AGCAGGATATTTGAGAGGAGATCTGAGTGGCTTGAAGGATGATGAGAGAGGAGACA 645  
Db 541 AGCAGGATATTTGAGAGGAGATCTGAGTGGCTTGAAGGATGATGAGAGAGGAGACA 600  
Qy 646 CAGGTGTGTCTTCCCTTGTGTATAGCGTGAAGGAGAGAGAGAGAGAGAGAGAGAG 704  
Db 601 CAGGTGTGTCTTCCCTTGTGTATAGCGTGAAGGAGAGAGAGAGAGAGAGAGAGAG 660  
Qy 705 TT-CCAAAGTGTGGTGGTCCGTTGGTCCGAGATACCTTTAGTGTATGGGAGCTG 763  
Db 661 TTCCAAAGTGTGGTGGTCCGTTGGTCCGAGATACCTTTAGTGTATGGGAGCTG 720  
Qy 764 CATTAAGTGCACAAATCAGAGCAAGAAAGCGATGCCCTCCCATTTCTCATCTCT 823  
Db 721 CATTAAGTGCACAAATCAGAGCAAGAAAGCGATGCCCTCCCATTTCTCATCTCT 780  
Qy 824 TTATGCCGAGAGATCTCAGCTGTGATGCCAATGTTCCGATGCTGTGAGAGCATGCC 883  
Db 781 TTATGCCGAGAGATCTCAGCTGTGATGCCAATGTTCCGATGCTGTGAGAGCATGCC 840  
Qy 884 GAGCTCTCTCTGCTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 943  
Db 841 GAGCTCTCTCTGCTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
Qy 944 CTTTATGAGACAGTGTGTGTATAGTGTATTTTATGCTACTTATGTTTATG 1003  
Db 901 CTTTATGAGACAGTGTGTGTATAGTGTATTTTATGCTACTTATGTTTATG 960  
Qy 1004 ACAAGTGTAGTGCATCTATTTATGACCTTTCAATTAATAGATTAAAGTAAAA 1063  
Db 961 ACAAGTGTAGTGCATCTATTTATGACCTTTCAATTAATAGATTAAAGTAAAA 1020  
Qy 1064 AAAAAAAAAAAAAA 1077  
Db 1021 AAAAAAAAAAAAAA 1034

RESULT 4  
US-09-882-171-58  
; Sequence 58, Application US/09882171  
; Publication No. US20030175858A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P2  
; CURRENT APPLICATION NUMBER: US/09/882,171  
; CURRENT FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 09/809,391  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 09/149,476  
; PRIOR FILING DATE: 1998-09-08  
; PRIOR APPLICATION NUMBER: PCT/US98/04493  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: 60/040,162  
; PRIOR FILING DATE: 1997-03-07



PRIOR APPLICATION NUMBER: 60/047,593  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,614  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/043,578  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/043,576  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/047,501  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/043,670  
PRIOR FILING DATE: 1997-04-11  
PRIOR APPLICATION NUMBER: 60/056,632  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,664  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,876  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,881  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,909  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,875  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,862  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,887  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/056,908  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/048,964  
PRIOR FILING DATE: 1997-06-06  
PRIOR APPLICATION NUMBER: 60/057,650  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: 60/056,884  
PRIOR FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: 60/057,669  
PRIOR FILING DATE: 1997-09-05

Query Match 66.4%; Score 792; DB 10; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

46 TCGGCGCTTGAGACAGACATCTACTAGGCTGGGTCACTCTGTGAATCACTGACTG 105  
1 TCGGCGCTTGAGACAGACATCTACTAGGCTGGGTCACTCTGTGAATCACTGACTG 60  
106 CAAAGCTCCCTCAATTTCTGTGAGCCATCAAGGACCCAGCCGCTTGGAGATGTT 165  
61 CAAAGCTCCCTCAATTTCTGTGAGCCATCAAGGACCCAGCCGCTTGGAGATGTT 120  
166 GCGGATCTTGGCCATGAGGAAATCTGTGAGAGAGACACCCCGAGTGAAGCCACTAC 225  
121 GCGGATCTTGGCCATGAGGAAATCTGTGAGAGAGACACCCCGAGTGAAGCCACTAC 180  
226 CCAAGCCCAAGAGGTAGCATTTCTTCAAGAGGCTTCTTCAATAGGGCCATGGTGTCC 285  
181 CCAAGCCCAAGAGGTAGCATTTCTTCAAGAGGCTTCTTCAATAGGGCCATGGTGTCC 240  
286 CCAAGGGGGTCTGAGCCCGCCAGAGAGAGAGGCTGAGGCTGGTGTCTGATGTC 345  
241 CCAAGGGGGTCTGAGCCCGCCAGAGAGAGAGGCTGAGGCTGGTGTCTGATGTC 300  
346 CCGCTTCAATACCTCAACCGGAGCTGTGTAATAGGGCTTTCGAGAGTGGCATCTCGG 405  
301 CCGCTTCAATACCTCAACCGGAGCTGTGTAATAGGGCTTTCGAGAGTGGCATCTCGG 360  
406 CAAAGCATCTGTGAGAGCGGTGACCTTCATCTGCTCTTCTGCTCATGATGTTGG 465  
361 CAAAGCATCTGTGAGAGCGGTGACCTTCATCTGCTCTTCTGCTCATGATGTTGG 420  
466 TGTTCGAGCTCTCTCTGTTGGCTTGTCTACTGTGTCCACCTGATTCAGCGGTG 525

421 TGTTCGAGCTCTCTCTGTTGGCTTGTCTACTGTGTCCACCTGATCAGCGGTG 480  
526 ACCTGTGAGGCTGTATAGGGGTGGGTTTGTGAGAGGACTTGTGGGCTTGTGTGAG 585  
481 ACCTGTGAGGCTGTATAGGGGTGGGTTTGTGAGAGGACTTGTGGGCTTGTGTGAG 540  
586 AGCAGCATATTTGAGAGGAGCATCTGTGTGGGCTTGAAGGATATATCAGAGAGGACCA 645  
541 AGCAGCATATTTGAGAGGAGCATCTGTGTGGGCTTGAAGGATATATCAGAGAGGACCA 600  
646 CAGGTGTGTGTTTCCCTTGTGTGAAGCGTGAAGGAGAGGAGAGCTTACT-CCAGCAT 704  
601 CAGGTGTGTGTTTCCCTTGTGTGAAGCGTGAAGGAGAGGAGAGCTTACT-CCAGCAT 660  
705 TT-CCAAAGTGTGGTGGTCCGTTGGTTCGCGAGATCTTTAGTGTGATGGGGCTG 763  
661 TTCCCAAGTGTGGTGGTCCGTTGGTTCGCGAGATCTTTAGTGTGATGGGGCTG 720  
764 CATTAGTGCAAAATCAGAGCAAGAAAGCATGCGCTTCCAAATCTCTCAATCCTT 823  
721 CATTAGTGCAAAATCAGAGCAAGAAAGCATGCGCTTCCAAATCTCTCAATCCTT 780  
824 TTATGCCGAGAAATCTGAGTGCATGATGCCAATGTTCCATGCTGTGAGAGACATGCC 883  
781 TTATGCCGAGAAATCTGAGTGCATGATGCCAATGTTCCATGCTGTGAGAGACATGCC 840  
884 GACGTCTCTCTGCTTAGGAGAGAGGACTTGGGCTTAGGAGAGGAGAGGAGGAGGAG 943  
841 GACGTCTCTCTGCTTAGGAGAGAGGACTTGGGCTTAGGAGAGGAGAGGAGGAGGAG 900  
944 CTTTGTAGCACTGTTTGTGTTTATGTAATATTTTATGCTACTTATGTTTGG 1003  
901 CTTTGTAGCACTGTTTGTGTTTATGTAATATTTTATGCTACTTATGTTTGG 960  
1004 ACAAGTGTAGTGAGGATCTTATTTATGTGACCTTTGAATTAATAGATTAGTAATA 1063  
961 ACAAGTGTAGTGAGGATCTTATTTATGTGACCTTTGAATTAATAGATTAGTAATA 1020

1064 AAAAAAAAAAAAAA 1077  
1021 AAAAAAAAAAAAAA 1034

RESULT 5  
US-10-164-861-58  
Sequence 58, Application US/10164861  
Publication No. US20030225248A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/10/164,861  
CURRENT FILING DATE: 2002-06-10  
PRIOR FILING DATE: 1998-09-08  
PRIOR APPLICATION NUMBER: US/09/149,476  
PRIOR FILING DATE: 1998-03-06  
NUMBER OF SEQ ID NOS: 757  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 58  
LENGTH: 1049  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-164-861-58  
Query Match 66.4%; Score 792; DB 13; Length 1049;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1032; Conservative 0; Mismatches 0; Indels 2; Gaps 2;







Db 275 GCCGAGAGATCTAGTGTGATGCAACATGTTCCGATGCTTGAGAAAGCATGCCGACG 334  
QY 888 TCTCCTCTGCTAGGAGCAGACTTGGGCTTAGGGCAGGTGAGAAAAATTCAGACTTT 947  
Db 335 TCTCCTCTGCTAGGAGCAGACTTGGGCTTAGGGCAGGTGAGAAAAATTCAGACTTT 394  
QY 948 TTTAG 952  
Db 395 TTTAG 399

RESULT 10  
US-09-864-761-5591  
Sequence 5591, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 5591  
LENGTH: 473  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC004514.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.3  
US-09-864-761-5591  
Query Match  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1048 TAGATTAAAGTAAAAA  
Db 259 TAGATTAAAGTAAAAA  
1085  
296

RESULT 11  
US-09-864-761-6261  
Sequence 6261, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 6261  
LENGTH: 473  
TYPE: DNA  
ORGANISM: Homo sapiens

OTHER INFORMATION: MAP TO AC004514.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.7  
US-09-864-761-6261

Query Match 3.2%; Score 38; DB 9; Length 473;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1048 TAGATTTAAGTAAAAA  
DB 259 TAGATTTAAGTAAAAA  
1085  
296

RESULT 12  
US-10-027-632-209023  
Sequence 209023, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 209023  
LENGTH: 998  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-209023

Query Match 3.2%; Score 38; DB 13; Length 998;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1048 TAGATTTAAGTAAAAA  
DB 307 TAGATTTAAGTAAAAA  
1085  
344

RESULT 13  
US-10-027-632-209023  
Sequence 209023, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 209023  
LENGTH: 998  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-209023

Query Match 3.2%; Score 38; DB 16; Length 998;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1048 TAGATTTAAGTAAAAA  
DB 307 TAGATTTAAGTAAAAA  
1085  
344

RESULT 14  
US-10-424-599-125925/C  
Sequence 125925, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 125925  
LENGTH: 616  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_84719C.1  
US-10-424-599-125925

Query Match 2.9%; Score 34; DB 13; Length 616;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1052 TTTAAGTAAAAA  
DB 65 TTTAAGTAAAAA  
1085  
32

RESULT 15  
US-10-678-521-24  
Sequence 521, Application US/10678521  
Publication No. US20040081998A1  
GENERAL INFORMATION:  
APPLICANT: Silver, Gary M.  
APPLICANT: Brandt, Kevin S.  
APPLICANT: Wisniewski, Nancy  
TITLE OF INVENTION: NOVEL CARBOXYL ESTERASE NUCLEIC ACID MOLECULES, PROTEINS AND USES  
FILE REFERENCE: FC-1-CI-PUS

Tue Aug 24 09:48:32 2004

us-10-001-885-31.011.rmpb

Page 10

; CURRENT APPLICATION NUMBER: US/10/678,521  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: US/09/403,942F  
; PRIOR FILING DATE: 2000-05-02  
; PRIOR APPLICATION NUMBER: PCT/US97/20598  
; PRIOR FILING DATE: 1997-11-10  
; PRIOR APPLICATION NUMBER: 08/747,221  
; PRIOR FILING DATE: 1996-11-12  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 2836  
; TYPE: DNA  
; ORGANISM: Ctenocephalides felis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (99)..(1889)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (2278)..(2278)  
; OTHER INFORMATION: n = unknown  
US-10-678-521-24

Query Match 2.98; Score 34; DB 17; Length 2836;  
Best Local Similarity 100.0%; Pred.No. 1.2e-06;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1052 TTTAAGTAAAAA  
Db 2778 TTTAAGTAAAAA  
1085  
2811

Search completed: August 22, 2004, 08:31:53  
Job time : 599 secs

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:17:04 ; Search time 54 Seconds

(without alignments)  
633.116 Million cell updates/sec

Title: US-10-001-885-125

Perfect score: 632  
Sequence: 1 MTRILANGRIYQDDPRVRT.....GYRGLLVGLVYVSHLSQR 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29aug04:\*  
1: Geneseq21980a:\*  
2: Geneseq21980a:\*  
3: Geneseq2000a:\*  
4: Geneseq2000a:\*  
5: Geneseq2002a:\*  
6: Geneseq2003a:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004a:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	632	100.0	121	AB079328	Ab079328 Human ova
2	632	100.0	121	AB061063	Ab061063 Human ova
3	568.5	90.0	120	AD060341	Ad060341 Rat Prote
4	568.5	90.0	120	AD060339	Ad060339 Rat Prote
5	387	61.2	111	AB035533	Ab035533 Human nov
6	387	61.2	111	AB034727	Ab034727 Fragment
7	245	38.8	48	AA074777	AA074777 Human sec
8	245	38.8	48	AB068995	Ab068995 Human pol
9	245	38.8	48	AB035227	Ab035227 Human nov
10	245	38.8	48	AB034421	Ab034421 Region of
11	184	29.1	114	AB030633	Ab030633 Novel hum
12	110.5	17.5	132	AB084669	Ab084669 Human SEC
13	110.5	17.5	132	AD080604	Ad080604 Novel pro
14	80	12.7	92	AA077595	AA077595 Human col
15	80	12.7	92	AA042404	AA042404 Human ORF
16	78.5	12.4	325	AA041419	AA041419 Paxillin
17	78.5	12.3	545	AA027493	AA027493 p34 from
18	77.5	12.3	1568	AB030109	Ab030109 Novel hum
19	75.5	11.9	622	AA068560	AA068560 Human nov
20	75.5	11.9	622	AB014288	Ab014288 Novel hum
21	75.5	11.9	622	AD033068	Ad033068 Human nov
22	75	11.9	417	AA012139	AA012139 Hydrophob
23	75	11.9	718	AA071059	AA071059 Human mem
24	75	11.9	718	AB052104	Ab052104 Homo sapi
25	75	11.9	735	AB035546	Ab035546 Human pro

#### ALIGNMENTS

RESULT 1	AB079328	standard; protein, 121 AA.
ID	AB079328	
AC	AB079328;	
DT	12-AUG-2002	(first entry)
DE	Human ovary specific protein SEQ ID NO:125.	
XX	Human; ovary specific nucleic acid; OSNA; ovary specific protein; OSP;	
KW	ovary specific gene; OSN; ovarian cancer; immune response; metastasis.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200240535-A2.	
XX		
PD	23-MAY-2002.	
XX		
PF	20-NOV-2001; 2001WO-US045011.	
XX		
PR	20-NOV-2000; 2000US-0252061P.	
PR	27-NOV-2000; 2000US-0253257P.	
PA	(DIAD-) DIADEXUS INC.	
XX		
PI	Salceda S, Macina RA, Recipon H, Caffertkey R, Sun Y, Liu C;	
XX		
DR	WPI; 2002-471617/50.	
XX		
PT	New ovary specific genes and proteins, useful as a vaccine for treating	
PT	patients with ovarian cancer, or for diagnosing and monitoring the	
PT	presence and metastases of ovarian cancer in a patient.	
XX		
PS	Claim 11; Page 241; 260pp; English.	
XX		
CC	AB087790 to AB087882 represent human ovary specific nucleic acid (OSNA)	
CC	sequences, and AB079297 to AB079370 represent ovary specific protein	
CC	(OSP) sequences from the present invention. OSNA and OSP sequences have	
CC	cytostatic activity, and can be used in vaccine production and gene	
CC	therapy. An antibody, that specifically binds to an OSP can be used for	
CC	treating a patient with ovarian cancer, particularly for inducing an	
CC	immune response against the ovarian cancer cell expressing the OSNA or	
CC	OSP. The OSNAs and OSPs can also be used for diagnosing and monitoring	
CC	the presence and metastases of ovarian cancer in a patient	
XX		
SO	Sequence 121 AA;	

26	73.5	11.6	442	4	AB0909021	Ab0909021 Novel hum
27	73	11.6	763	3	AA042392	AA042392 Arabidops
28	73	11.6	793	3	AA042391	AA042391 Arabidops
29	72.5	11.5	315	4	AB061702	Ab061702 Drosophi
30	72	11.3	310	2	AA023996	AA023996 B-HDAg-T
31	71.5	11.3	158	3	AA022606	AA022606 Zea mays
32	71.5	11.3	19938	6	AB076680	Ab076680 Streptomy
33	71	11.2	124	4	AA001631	AA001631 Human pol
34	71	11.2	345	3	AA093600	AA093600 Protein e
35	71	11.2	349	4	AB04137	Ab04137 Novel hum
36	71	11.2	689	2	AA072025	AA072025 HSV-2 str
37	71	11.2	1030	4	AB022699	Ab022699 Novel hum
38	71	11.2	1697	7	AD037542	Ad037542 Human nuc
39	70.5	11.2	226	3	AA042794	AA042794 Human CRF
40	70.5	11.2	915	5	AB063120	Ab063120 FLO11 gen
41	70	11.1	213	6	AB073050	Ab073050 Amino aci
42	70	11.1	266	4	AB018528	Ab018528 Novel hum
43	70	11.1	379	7	AD085283	Ad085283 Rat foocce
44	70	11.1	387	4	AB011608	Ab011608 Novel hum
45	70	11.1	1128	4	AB058911	Ab058911 Drosophi

Query Match 100.0%; Score 632; DB 5; Length 121;  
Best Local Similarity 100.0%; Pred. No. 8.9e-62;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRILANGELIVODDDPRVRRITTOPPGRSIPROSPFNRGHGAPGPGPQQOAGARLGA 60  
DB 1 MVRILANGELIVODDDPRVRRITTOPPGRSIPROSPFNRGHGAPGPGPQQOAGARLGA 60  
QY 61 QSPFNDLNRQLVNNGFPQWHLGNHAEVPVTSILLLFLMMLGVRGILLVGLVYLVSHLSQ 120  
DB 61 QSPFNDLNRQLVNNGFPQWHLGNHAEVPVTSILLLFLMMLGVRGILLVGLVYLVSHLSQ 120  
QY 121 R 121  
DB 121 R 121

## RESULT 2

ABU61063 standard; protein; 121 AA.

ABU61063;  
ABU61063;  
08-MAY-2003 (first entry)  
Human ovarian specific protein DEX0310\_122.

Human; ovarian specific protein; OSP; ovarian cancer;  
non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;  
gynaecological.

Homo sapiens.

WO200292785-A2.

21-NOV-2002.

13-FEB-2002; 2002MO-US022271.

13-FEB-2001; 2001US-0268290P.

15-FEB-2001; 2001US-0268834P.

(DIAD-) DIADEXUS INC.

Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferty R;  
Sun Y, Liu C;

WPI; 2003-120677/11.

N-PSDB; ABX92251.

New isolated OSGNA nucleic acid and encoded polypeptide, useful for  
identifying, diagnosing, monitoring, staging, imaging and treating  
ovarian cancer and non-cancerous diseases in ovarian tissues.

Claim 11; Page 220-221; 224pp; English.

The invention relates to a new isolated nucleic acid termed ovarian  
specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that

encodes any of 53 fully defined protein sequences appearing as ABU61018-

ABU61070 (termed ovarian specific proteins, OSP); (b) any of 76 fully

defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a

molecule of (a) or (b). Also included are a method for determining the

presence of an OSGNA, a host cell comprising the vector, an isolated OSP

comprising an OSGNA, an anti-OSP antibody or fragment, a method for determining

the presence of an OSGNA, the methods and compositions of the present

invention are useful for identifying, diagnosing, monitoring, staging,  
imaging and treating ovarian cancer and non-cancerous disease in ovary  
tissue. The present sequence is an OSP of the invention

Query Match 100.0%; Score 632; DB 6; Length 121;  
Best Local Similarity 100.0%; Pred. No. 8.9e-62;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRILANGELIVODDDPRVRRITTOPPGRSIPROSPFNRGHGAPGPGPQQOAGARLGA 60  
DB 1 MVRILANGELIVODDDPRVRRITTOPPGRSIPROSPFNRGHGAPGPGPQQOAGARLGA 60  
QY 61 QSPFNDLNRQLVNNGFPQWHLGNHAEVPVTSILLLFLMMLGVRGILLVGLVYLVSHLSQ 120  
DB 61 QSPFNDLNRQLVNNGFPQWHLGNHAEVPVTSILLLFLMMLGVRGILLVGLVYLVSHLSQ 120  
QY 121 R 121  
DB 121 R 121

## RESULT 3

ADE60341 standard; protein; 120 AA.

ADE60341;  
ADE60341;  
29-JAN-2004 (first entry)

Rat Protein BAB25613, SEQ ID NO 6247.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

Rattus norvegicus.

WO2003016475-A2.

27-FEB-2003.

14-AUG-2002; 2002MO-US025765.

14-AUG-2001; 2001US-0312147P.

01-NOV-2001; 2001US-0346382P.

26-NOV-2001; 2001US-0333347P.

(GENO) GEN HOSPITAL CORP.

(FARB) BAYER AG.

Woolf C, D'Urso D, Befort K, Costigan M;  
WPI; 2003-268312/26.

GENBANK; BAB25613.

New composition comprising two or more isolated polypeptides, useful for  
preparing a medicament for treating pain in an animal.

Claim 1; Page: 1017pp; English.

The invention discloses a composition comprising two or more isolated rat  
or human polynucleotides or a polynucleotide which represents a fragment,

derivative or allelic variation of the nucleic acid sequence. Also  
claimed are a vector comprising the novel polynucleotide, a host cell

comprising the vector, a method for identifying a nucleotide sequence  
which is differentially regulated in an animal subjected to pain and a

kit to perform the method, an array, a method for identifying an agent  
that increases or decreases the expression of the polynucleotide sequence

that is differentially expressed in neuronal tissue of a first animal  
subjected to pain, a method for identifying a compound which regulates

the expression of a polynucleotide sequence which is differentially  
expressed in an animal subjected to pain, a method for identifying a

compound that regulates the activity of one or more of the  
polynucleotides, a method for producing a pharmaceutical composition, a

method for identifying a compound or small molecule that regulates the  
activity in an animal of one or more of the polypeptides given in the  
specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 120 AA:

Query Match 90.0%; Score 568.5; DB 7; Length 120;

Best Local Similarity 92.6%; Pred. No. 9.1e-55;

Matches 112; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MRLIANGELIVDDPRVETTTQPPRGSIHQSFENRGGAPPGGPRQQAARLGAA 60

DB 1 MRLIANGELIVDDPRVETTTQ-HRSSSSQCGFNRGGAPPGGPRQQAARLGAA 59

QY 61 QSPFDLNRQLVNMGGFPQWHLGNHVEPTVTSILLFLMLGVGGLLVGLVIVSHLSQ 120

DB 60 QSPFDLNRQLVNMGGFPQWHLGNHVEPTVTSILLFLMLGVGGLLVGLVIVSHLSQ 119

QY 121 R 121

DB 120 R 120

RESULT 4  
 ADE60339 standard; protein; 120 AA.

AC ADE60339;

DT 29-JAN-2004 (first entry)

DE Rat Protein BAB5613, SEQ ID NO 6245.

KM Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

OS Rattus norvegicus.

PN MO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WC-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

PA (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI; 2003-268312/26.

DR GENBANK; BAB5613.

PT New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017P; English.

XX The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotide or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also cell

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 120 AA:

Query Match 90.0%; Score 568.5; DB 7; Length 120;

Best Local Similarity 92.6%; Pred. No. 9.1e-55;

Matches 112; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 MRLIANGELIVDDPRVETTTQPPRGSIHQSFENRGGAPPGGPRQQAARLGAA 60

DB 1 MRLIANGELIVDDPRVETTTQ-HRSSSSQCGFNRGGAPPGGPRQQAARLGAA 59

QY 61 QSPFDLNRQLVNMGGFPQWHLGNHVEPTVTSILLFLMLGVGGLLVGLVIVSHLSQ 120

DB 60 QSPFDLNRQLVNMGGFPQWHLGNHVEPTVTSILLFLMLGVGGLLVGLVIVSHLSQ 119

QY 121 R 121

DB 120 R 120

RESULT 5  
 ABG95533 standard; protein; 111 AA.

AC ABG95533;

DT 15-JAN-2003 (first entry)

DE Human novel secreted protein gene 48 polypeptide #1.

KM Human; secreted protein; autoimmune disease; chemotaxis;

KW rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;

KW liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;

KW cerebrovascular disorder; cerebral ischemia; angiogenesis; sunburn;

KW nervous system disorders; Alzheimer's disease; infection;

KW ocular disorder; corneal infection; wound healing; tissue regeneration;

KW epithelial cell proliferation; organ transplantation; food additive;

XX preservative; nutritional.

OS Homo sapiens.

PN US6420526-B1.

PD 16-JUL-2002.

PF 08-SEP-1998; 98US-00149476.

PR 07-MAR-1997; 97US-0038623P.

PR 07-MAR-1997; 97US-0040161P.

PR 07-MAR-1997; 97US-0040162P.

PR 07-MAR-1997; 97US-0040163P.





Db 99 QSPFNRLNRLVNV 111

RESULT 6  
ABO34727  
ID ABO34727 standard; protein; 111 AA.  
XX ABO34727;  
AC  
XX  
XX  
DT 22-SEP-2003 (first entry)  
DE  
XX  
XX  
Fragment #45 of a human secreted protein.

Human; secreted protein; hyperproliferative disorder; leukaemia;  
breast cancer; wound; reproductive disorder; blood-related disorder;  
haemophilia; thrombocytopaenia; immunodeficiency; chronic hypoplasia;  
Wiskott-Aldrich syndrome; autoimmune disorder; multiple sclerosis;  
graft-versus-host disease; Hashimoto's thyroiditis; allergy; asthma;  
viral infection; bacterial infection; fungal infection; AIDS; sepsis;  
renal disorder; kidney failure; cardiovascular disorder; cytostatic;  
angina pectoris; cerebral ischaemia; congenital heart defect;  
respiratory disorder; neurological disorder; Alzheimer's disease;  
Parkinson's disease; inflammation; Crohn's disease; vulnery;  
immunosuppressive; antibacterial; haemostatic; thrombolytic;  
anticoagulant; neuroprotective; thymolytic; antiallergic;  
antidiabetic; virucide; fungicide; anti-HIV; nephrotoxic; antiangiinal;  
cerebroprotective; cardiant; nootropic; antiparkinsonian;  
antiinflammatory.

XX Homo sapiens.  
XX  
XX US2003049618-A1.  
XX  
XX 13-MAR-2003.  
XX  
XX 16-MAR-2001; 2001US-00809391.  
XX  
XX 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040335P.  
PR 07-MAR-1997; 97US-0040626P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-0043659P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 23-MAY-1997; 97US-0043674P.  
PR 23-MAY-1997; 97US-0047482P.  
PR 23-MAY-1997; 97US-0047500P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047502P.  
PR 23-MAY-1997; 97US-0047503P.  
PR 23-MAY-1997; 97US-0047504P.  
PR 23-MAY-1997; 97US-0047582P.  
PR 23-MAY-1997; 97US-0047583P.  
PR 23-MAY-1997; 97US-0047584P.  
PR 23-MAY-1997; 97US-0047585P.  
PR 23-MAY-1997; 97US-0047586P.  
PR 23-MAY-1997; 97US-0047587P.  
PR 23-MAY-1997; 97US-0047588P.  
PR 23-MAY-1997; 97US-0047589P.

PR 23-MAY-1997; 97US-0047590P.  
PR 23-MAY-1997; 97US-0047592P.  
PR 23-MAY-1997; 97US-0047593P.  
PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047616P.  
PR 23-MAY-1997; 97US-0047617P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047632P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 06-JUN-1997; 97US-0048974P.  
PR 13-JUN-1997; 97US-0049610P.  
PR 08-JUL-1997; 97US-0051926P.  
PR 16-JUL-1997; 97US-0052872P.  
PR 16-AUG-1997; 97US-005724P.  
PR 22-AUG-1997; 97US-005630P.  
PR 22-AUG-1997; 97US-005631P.  
PR 22-AUG-1997; 97US-005632P.  
PR 22-AUG-1997; 97US-005633P.  
PR 22-AUG-1997; 97US-005636P.  
PR 22-AUG-1997; 97US-005637P.  
PR 22-AUG-1997; 97US-005662P.  
PR 22-AUG-1997; 97US-005664P.  
PR 22-AUG-1997; 97US-0056645P.  
PR 22-AUG-1997; 97US-0056862P.  
PR 22-AUG-1997; 97US-0056864P.  
PR 22-AUG-1997; 97US-0056872P.  
PR 22-AUG-1997; 97US-0056874P.  
PR 22-AUG-1997; 97US-0056875P.  
PR 22-AUG-1997; 97US-0056876P.  
PR 22-AUG-1997; 97US-0056877P.  
PR 22-AUG-1997; 97US-0056878P.  
PR 22-AUG-1997; 97US-0056879P.  
PR 22-AUG-1997; 97US-0056880P.  
PR 22-AUG-1997; 97US-0056881P.  
PR 22-AUG-1997; 97US-0056882P.  
PR 22-AUG-1997; 97US-0056884P.  
PR 22-AUG-1997; 97US-0056886P.  
PR 22-AUG-1997; 97US-0056887P.  
PR 22-AUG-1997; 97US-0056888P.  
PR 22-AUG-1997; 97US-0056889P.  
PR 22-AUG-1997; 97US-0056892P.  
PR 22-AUG-1997; 97US-0056893P.  
PR 22-AUG-1997; 97US-0056894P.  
PR 22-AUG-1997; 97US-0056903P.  
PR 22-AUG-1997; 97US-0056908P.  
PR 22-AUG-1997; 97US-0056909P.  
PR 22-AUG-1997; 97US-0056910P.  
PR 22-AUG-1997; 97US-0056911P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057659P.  
PR 12-SEP-1997; 97US-0057761P.  
PR 12-SEP-1997; 97US-0058785P.  
PR 09-OCT-1997; 97US-0061660P.  
PR 06-MAR-1998; 98WO-US004493.  
PR 08-SEP-1998; 98US-0014947C.  
PR 17-MAR-2000; 2000US-0190068P.

XX  
XX (RUBE/) RUBEN S. M.  
PA (ROSE/) ROSEN C. A.  
PA (SOPP/) SOPPET D. R.  
PA (CARP/) CARTER K. C.  
PA (BEDN/) BEDNARIK D. P.  
PA (ENDR/) ENDRESS G. A.

PA (YUGG/) YU G.  
PA (NIJ/) NI J.  
PA (FENG/) FENG P.  
PA (YOUN/) YOUNG P. E.  
PA (GREEN/) GREENE J. M.  
PA (FERRI/) FERRIE A. M.  
PA (DUAN/) DUAN D. R.  
PA (HUJ/) HU J.  
PA (FLORE/) FLORENCE K. A.  
PA (OLSE/) OLSEN H. S.  
PA (FISC/) FISCHER C. L.  
PA (EBNER/) EBNER R.  
PA (BREW/) BREWER L. A.  
PA (MOOR/) MOORE P. A.  
PA (SHIY/) SHI Y.  
PA (LAFLE/) LAFLEUR D. W.  
PA (LIY/) LI Y.  
PA (ZENG/) ZENG Z.  
PA (KJAW/) KJAW H.  
XX  
XX Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP, Fendres GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Ferrie AM, Duan DR, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R, Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H; WPI; 2003-521800/49.  
XX  
XX New genes and its encoded prostate cancer antigen proteins, useful for preventing, treating, ameliorating or diagnosing e.g. prostate cancers, lymphic hypoplasia, multiple sclerosis, AIDS, angina pectoris or cerebral ischemia.  
XX  
XX Claim 3; Page 29; 260pp; English.  
XX  
XX The present invention relates to the isolation of novel human secreted proteins and the polynucleotide sequences encoding them. The invention also discloses vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The polypeptide and polynucleotide sequences for the secreted proteins are useful for preventing, treating, ameliorating or diagnosing medical conditions such as hyperproliferative disorders (e.g. leukemia or breast cancers), wounds, reproductive disorders, blood-related disorders (e.g. haemophilia or thrombocytopenia), immunodeficiencies (e.g. Wiskott-Aldrich syndrome or thymic hypoplasia), autoimmune disorders (e.g. graft-versus-host disease, multiple sclerosis or Hashimoto's thyroiditis), allergies (e.g. asthma), viral or bacterial or fungal infections (e.g. AIDS or sepsis), renal disorders (e.g. kidney failure), cardiovascular disorders (e.g. angina pectoris, cerebral ischaemia or congenital heart defects), respiratory disorders, neurological disorders (e.g. Alzheimer's disease or Parkinson's disease), and inflammations (e.g. Crohn's disease). The polynucleotide or polypeptide may also be used as vaccine adjuvants.  
XX AB034374-AB034815 represent human secreted proteins or their fragments.  
XX Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO web site at [seqdata.uspto.gov/patseq/identity.html](http://seqdata.uspto.gov/patseq/identity.html)  
XX  
XX Sequence 111 AA;

Query Match 61.2%; Score 387; DB 6; Length 111;  
Best Local Similarity 100.0%; Freq. No. 9,5e-35;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVRILANGELVODDDPRVTTTTPPGSGIRPGSFFNRGHCAPGCGPRQCGARLGA 60  
Db 39 MVRILANGELVODDDPRVTTTTPPGSGIRPGSFFNRGHCAPGCGPRQCGARLGA 98  
Qy 61 QSPENDINROLVN 73  
Db 99 QSPENDINROLVN 111

RESULT 7  
AAW74777

ID AAW74777 standard; protein; 48 AA.  
XX  
AC AAW74777;  
XX  
DT 19-JAN-1999 (first entry)  
XX  
DE Human secreted protein encoded by gene 48 clone HFCAT74.  
XX  
XX Human; secreted protein; testis; tumour; foetal brain tissue;  
XX fusion protein; cancer; central nervous system; seizure; diagnosis;  
XX neurodegenerative disease.  
XX  
XX Homo sapiens.  
XX  
XX MO9839448-A2.  
XX  
PD 11-SEP-1998.  
XX  
PF 06-MAR-1998; 98MO-US004493.  
XX  
XX 07-MAR-1997; 97US-0038621P.  
XX 07-MAR-1997; 97US-0040161P.  
XX 07-MAR-1997; 97US-0040162P.  
XX 07-MAR-1997; 97US-0040163P.  
XX 07-MAR-1997; 97US-0040333P.  
XX 07-MAR-1997; 97US-0040334P.  
XX 07-MAR-1997; 97US-0040335P.  
XX 07-MAR-1997; 97US-0040336P.  
XX 07-MAR-1997; 97US-0040337P.  
XX 07-MAR-1997; 97US-0040338P.  
XX 07-MAR-1997; 97US-0040339P.  
XX 07-MAR-1997; 97US-0040340P.  
XX 07-MAR-1997; 97US-0040341P.  
XX 07-MAR-1997; 97US-0040342P.  
XX 07-MAR-1997; 97US-0040343P.  
XX 07-MAR-1997; 97US-0040344P.  
XX 07-MAR-1997; 97US-0040345P.  
XX 07-MAR-1997; 97US-0040346P.  
XX 07-MAR-1997; 97US-0040347P.  
XX 07-MAR-1997; 97US-0040348P.  
XX 07-MAR-1997; 97US-0040349P.  
XX 07-MAR-1997; 97US-0040350P.  
XX 07-MAR-1997; 97US-0040351P.  
XX 07-MAR-1997; 97US-0040352P.  
XX 07-MAR-1997; 97US-0040353P.  
XX 07-MAR-1997; 97US-0040354P.  
XX 07-MAR-1997; 97US-0040355P.  
XX 07-MAR-1997; 97US-0040356P.  
XX 07-MAR-1997; 97US-0040357P.  
XX 07-MAR-1997; 97US-0040358P.  
XX 07-MAR-1997; 97US-0040359P.  
XX 07-MAR-1997; 97US-0040360P.  
XX 07-MAR-1997; 97US-0040361P.  
XX 07-MAR-1997; 97US-0040362P.  
XX 07-MAR-1997; 97US-0040363P.  
XX 07-MAR-1997; 97US-0040364P.  
XX 07-MAR-1997; 97US-0040365P.  
XX 07-MAR-1997; 97US-0040366P.  
XX 07-MAR-1997; 97US-0040367P.  
XX 07-MAR-1997; 97US-0040368P.  
XX 07-MAR-1997; 97US-0040369P.  
XX 07-MAR-1997; 97US-0040370P.  
XX 07-MAR-1997; 97US-0040371P.  
XX 07-MAR-1997; 97US-0040372P.  
XX 07-MAR-1997; 97US-0040373P.  
XX 07-MAR-1997; 97US-0040374P.  
XX 07-MAR-1997; 97US-0040375P.  
XX 07-MAR-1997; 97US-0040376P.  
XX 07-MAR-1997; 97US-0040377P.  
XX 07-MAR-1997; 97US-0040378P.  
XX 07-MAR-1997; 97US-0040379P.  
XX 07-MAR-1997; 97US-0040380P.  
XX 07-MAR-1997; 97US-0040381P.  
XX 07-MAR-1997; 97US-0040382P.  
XX 07-MAR-1997; 97US-0040383P.  
XX 07-MAR-1997; 97US-0040384P.  
XX 07-MAR-1997; 97US-0040385P.  
XX 07-MAR-1997; 97US-0040386P.  
XX 07-MAR-1997; 97US-0040387P.  
XX 07-MAR-1997; 97US-0040388P.  
XX 07-MAR-1997; 97US-0040389P.  
XX 07-MAR-1997; 97US-0040390P.  
XX 07-MAR-1997; 97US-0040391P.  
XX 07-MAR-1997; 97US-0040392P.  
XX 07-MAR-1997; 97US-0040393P.  
XX 07-MAR-1997; 97US-0040394P.  
XX 07-MAR-1997; 97US-0040395P.  
XX 07-MAR-1997; 97US-0040396P.  
XX 07-MAR-1997; 97US-0040397P.  
XX 07-MAR-1997; 97US-0040398P.  
XX 07-MAR-1997; 97US-0040399P.  
XX 07-MAR-1997; 97US-0040400P.  
XX 07-MAR-1997; 97US-0040401P.  
XX 07-MAR-1997; 97US-0040402P.  
XX 07-MAR-1997; 97US-0040403P.  
XX 07-MAR-1997; 97US-0040404P.  
XX 07-MAR-1997; 97US-0040405P.  
XX 07-MAR-1997; 97US-0040406P.  
XX 07-MAR-1997; 97US-0040407P.  
XX 07-MAR-1997; 97US-0040408P.  
XX 07-MAR-1997; 97US-0040409P.  
XX 07-MAR-1997; 97US-0040410P.  
XX 07-MAR-1997; 97US-0040411P.  
XX 07-MAR-1997; 97US-0040412P.  
XX 07-MAR-1997; 97US-0040413P.  
XX 07-MAR-1997; 97US-0040414P.  
XX 07-MAR-1997; 97US-0040415P.  
XX 07-MAR-1997; 97US-0040416P.  
XX 07-MAR-1997; 97US-0040417P.

PR 23-MAY-1997; 97US-0047618P.  
 PR 23-MAY-1997; 97US-0047632P.  
 PR 23-MAY-1997; 97US-0047633P.  
 PR 06-JUN-1997; 97US-0048964P.  
 PR 06-JUN-1997; 97US-0048974P.  
 PR 13-JUN-1997; 97US-0049610P.  
 PR 08-JUL-1997; 97US-0051926P.  
 PR 16-JUL-1997; 97US-0052874P.  
 PR 18-AUG-1997; 97US-0055724P.  
 PR 22-AUG-1997; 97US-0056630P.  
 PR 22-AUG-1997; 97US-0056631P.  
 PR 22-AUG-1997; 97US-0056632P.  
 PR 22-AUG-1997; 97US-0056636P.  
 PR 22-AUG-1997; 97US-0056637P.  
 PR 22-AUG-1997; 97US-0056662P.  
 PR 22-AUG-1997; 97US-0056664P.  
 PR 22-AUG-1997; 97US-0056845P.  
 PR 22-AUG-1997; 97US-0056862P.  
 PR 22-AUG-1997; 97US-0056864P.  
 PR 22-AUG-1997; 97US-0056872P.  
 PR 22-AUG-1997; 97US-0056874P.  
 PR 22-AUG-1997; 97US-0056875P.  
 PR 22-AUG-1997; 97US-0056876P.  
 PR 22-AUG-1997; 97US-0056877P.  
 PR 22-AUG-1997; 97US-0056878P.  
 PR 22-AUG-1997; 97US-0056879P.  
 PR 22-AUG-1997; 97US-0056880P.  
 PR 22-AUG-1997; 97US-0056881P.  
 PR 22-AUG-1997; 97US-0056882P.  
 PR 22-AUG-1997; 97US-0056884P.  
 PR 22-AUG-1997; 97US-0056886P.  
 PR 22-AUG-1997; 97US-0056887P.  
 PR 22-AUG-1997; 97US-0056888P.  
 PR 22-AUG-1997; 97US-0056889P.  
 PR 22-AUG-1997; 97US-0056892P.  
 PR 22-AUG-1997; 97US-0056893P.  
 PR 22-AUG-1997; 97US-0056894P.  
 PR 22-AUG-1997; 97US-0056903P.  
 PR 22-AUG-1997; 97US-0056908P.  
 PR 22-AUG-1997; 97US-0056909P.  
 PR 22-AUG-1997; 97US-0056910P.  
 PR 22-AUG-1997; 97US-0056911P.  
 PR 05-SEP-1997; 97US-0057650P.  
 PR 05-SEP-1997; 97US-0057659P.  
 PR 05-SEP-1997; 97US-0057761P.  
 PR 12-SEP-1997; 97US-0058785P.  
 PR 02-OCT-1997; 97US-0061060P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Rosen CA, Fischer CL, Soppet DR, Carter KC,  
 PI Bednarik DP, Endress GA, Yu G, Ni J, Feng P, Young PS, Greene JM,  
 PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA,  
 PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;  
 XX WPI; 1998-506364/43.  
 DR N-PSDB; AAV59558.  
 DR  
 XX  
 XX New isolated human genes and the secreted polypeptide (s) they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders.  
 XX  
 XX Claim 1; Page 567; 721pp; English.  
 PS  
 XX This sequence represents a secreted human protein encoded by the nucleic  
 CC acid molecule designated Gene 48 from the human cDNA clone HFCAT174  
 CC (deposited as clone ATCC 97899 and ATCC 209045). The gene can be used to  
 CC generate fusion proteins by linking to the gene to a human immunoglobulin  
 CC portion (e.g. AAV59502) for increasing the stability of the fused  
 CC protein as compared to the human protein only. The invention relates to  
 CC 186 novel genes and their fragments (nucleic acid sequences: AAV5511-  
 CC V5812; amino acid sequences AAV74731-W75026) which are useful for  
 CC preventing, treating or ameliorating medical conditions e.g. by protein

CC or gene therapy. Also, pathological conditions can be diagnosed by  
 CC determining the amount of the new polypeptides in a sample or by  
 CC determining the presence of mutations in the new polynucleotides.  
 CC Specific uses are described for each of the 186 polynucleotides, based on  
 CC which tissues they are most highly expressed in (see AAV59511 for  
 CC described uses)  
 XX  
 XX Sequence 48 AA;  
 SQ  
 Query Match 38.8%; Score 245; DB 2; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 1,7e-19;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 74 MGFPQWHLGNHVAPEVTSILLLFILMLGVKGLLVGLVYVSHLSQR 121  
 DB 1 MGFPQWHLGNHVAPEVTSILLLFILMLGVKGLLVGLVYVSHLSQR 48  
 RESULT 8  
 ABP68995  
 ID ABP68995 standard; protein; 48 AA.  
 XX  
 XX ABP68995;  
 AC  
 XX 20-JAN-2003 (first entry)  
 DT  
 XX  
 XX Human polypeptide SEQ ID NO 1042.  
 DE  
 XX  
 XX Human; genome mapping; gene therapy; food supplement; virus; fungus;  
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;  
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
 KW arthritis; cytostatic; immunomodulator; nociceptive; neuroprotective;  
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
 KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
 KW antiaerthritic.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO200270539-A2.  
 PN  
 XX  
 XX 12-SEP-2002.  
 PD  
 XX  
 XX 05-MAR-2002; 2002WO-US005095.  
 PF  
 XX  
 XX 05-MAR-2001; 2001US-00799451.  
 PR  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F,  
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
 PI Wehrman T, Wang J, Wang D, Drmanac RT;  
 XX WPI; 2002-759812/82.  
 DR N-PSDB; AB211212.  
 DR  
 XX  
 XX New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative,  
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
 PT or coagulation disorders.  
 XX  
 XX Claim 9; SEQ ID NO 1042; 1012pp + Sequence listing; English.  
 PS  
 XX  
 XX The invention relates to an isolated polynucleotide (1) comprising a  
 CC nucleotide sequence selected from any of 948 sequences (AB21119-  
 CC AB21066) or their mature protein coding portion, active domain coding  
 CC protein or complementary sequences. The polynucleotides are useful for  
 CC identifying expressed genes or for physical mapping of human genome. The  
 CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
 CC markers, as a food supplement, for generating antibodies, in medical  
 CC imaging, screening and diagnostic assays and for treating cell-  
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,



XX (HUMA-) HUMAN GENOME SCI INC.  
XX Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC;  
PI Bednarik DR, Endress GA, Yu G, Ni U, Feng P, Young PE, Greene JM;  
PI Ferrite AM, Duan R, Hu J, Florence KA, Olsen HS, Ehner R, Brewer LA;  
PI Moore PA, Shi Y, Latleur DM, Li Y, Zeng Z, Kyaw H;  
XX MPI; 2002-634796/68.  
DR N-PSDB; ABS73545.  
XX  
PT New isolated human secreted protein for diagnosing, preventing, treating  
PT or ameliorating medical conditions and used as a food additive or  
PT preservative.  
XX  
PS Example 1; SEQ ID NO 367; 129p; English.  
XX  
CC The invention relates to an isolated protein that is one of 186 human  
CC secreted proteins, given in the specification, encoded by one of 309 cDNA  
CC sequences also given in the specification. The protein is used in a  
CC pharmaceutical composition used to prevent, treat or ameliorate a medical  
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. Disorders which are diagnosed or treated include  
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC e.g. corneal infection. The polypeptides can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,  
CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
CC components. The present sequence represents one of the novel human  
CC secreted proteins of the invention. Note: This sequence did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from USPTO at seqdata.uspto.gov/sequence.htmlDocID=6420526B1  
XX  
SQ Sequence 48 AA;  
Query Match 38.8%; Score 245; DB 5; Length 48;  
Best Local Similarity 100.0%; Pred. No. 1.7e-19;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 74 MGFPQWHLGNHAYEPVTSILLFLMMLGVRGLLVGLVYLVHLSQR 121  
Db 1 MGFPQWHLGNHAYEPVTSILLFLMMLGVRGLLVGLVYLVHLSQR 48  
RESULT 10  
ID ABO34421 standard; protein; 48 AA.  
XX  
AC ABO34421;  
XX  
DT 22-SEP-2003 (first entry)  
XX  
DE Region of human secreted protein encoded by cDNA sequence #48.  
XX  
KM Human; secreted protein; hyperproliferative disorder; leukaemia;  
KM breast cancer; wound; reproductive disorder; blood-related disorder;  
KM hemophilia; thrombocytopenia; immunodeficiency; thymic hypoplasia;  
KM Wiskott-Aldrich syndrome; autoimmune disorder; multiple sclerosis;  
KM graft-versus-host disease; Hashimoto's thyroiditis; allergy; asthma;  
KM viral infection; bacterial infection; fungal infection; AIDS; sepsis;  
KM renal disorder; kidney failure; cardiovascular disorder; cytostatic;  
KM angina pectoris; cerebral ischaemia; congenital heart defect;  
KM respiratory disorder; neurological disorder; Alzheimer's disease;  
KM Parkinson's disease; inflammation; Crohn's disease; vulnerability;  
KM immunosuppressive; antibacterial; haemostatic; thrombolytic;  
KM anticoagulant; neuroprotective; thyromimetic; antiallergic;

KM antiasthmatic; virucide; fungicide; anti-HIV; nephrotoxic; antianginal;  
KM cerebroprotective; cardiant; nootropic; antiparkinsonian;  
KM antiinflammatory.  
XX  
OS Homo sapiens.  
XX  
PN US2003049618-A1.  
XX  
PD 13-MAR-2003.  
XX  
PF 16-MAR-2001; 2001US-00809391.  
XX  
PR 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040626P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-0043669P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043674P.  
PR 23-MAY-1997; 97US-0047492P.  
PR 23-MAY-1997; 97US-0047500P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047502P.  
PR 23-MAY-1997; 97US-0047503P.  
PR 23-MAY-1997; 97US-0047581P.  
PR 23-MAY-1997; 97US-0047582P.  
PR 23-MAY-1997; 97US-0047583P.  
PR 23-MAY-1997; 97US-0047584P.  
PR 23-MAY-1997; 97US-0047585P.  
PR 23-MAY-1997; 97US-0047586P.  
PR 23-MAY-1997; 97US-0047587P.  
PR 23-MAY-1997; 97US-0047588P.  
PR 23-MAY-1997; 97US-0047589P.  
PR 23-MAY-1997; 97US-0047590P.  
PR 23-MAY-1997; 97US-0047592P.  
PR 23-MAY-1997; 97US-0047593P.  
PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047617P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047632P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 06-JUN-1997; 97US-0048974P.  
PR 13-JUN-1997; 97US-0049610P.  
PR 08-JUL-1997; 97US-0051926P.  
PR 16-JUL-1997; 97US-0052874P.  
PR 18-AUG-1997; 97US-0055724P.



XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 20; SEQ ID NO 30992; 103bp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (II) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantifying a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pcr\_sequences  
CC  
XX Sequence 114 AA;  
SQ  
Query Match 29.1%; Score 184; DB 4; Length 114;  
Best Local Similarity 44.4%; Pred. No. 2.6e-12;  
Matches 56; Conservative 4; Mismatches 38; Indels 28; Gaps 6;  
QY 1 MVRLLANGELVODDDPRVTRTTQPPRSGIFRNGHAPGPGPPROQAGAR-LGA 59  
DB 1 MVRLLANGELVODDRP-----PSEDRP-----ATKRSVAAPAKQRCQGMVL 44  
QY 60 AQSEFNDLRLQ-LVNMGPQWHLGNHAYEPY-----TSILLFLMMLGVRLILVGLV 112  
DB 45 LSSFNDLKPAAGRTWAFRSWHLR-----QPLVAGSDPSCILFLMMLGVRLILVGLV 100  
QY 113 YLVSHL 118  
DB 101 YLVSHL 106  
RESULT 12  
ABBS4669  
ID ABBS4669 standard; protein; 132 AA.  
XX  
XX ABB84669;  
AC  
XX  
XX 11-FEB-2003 (first entry)  
DT  
XX  
XX Human SECP-21 protein from clone 1265837CD1 SEQ ID 21.  
DE  
XX  
XX Secreted protein; SECP; human; antiarteriosclerotic; antiatherosclerotic;  
KM hepatotropic; cytosarctic; anti-HIV; antiallergic; antiasthmatic; cancer;  
KM antianaemic; antidiabetic; antiinflammatory; neuroprotective; anticancer;  
KM antineumatic; antiarthritic; cardiatic; hypotensive; gonadal dysgenesis;  
KM antiparkinsonian; anticonvulsant; nootropic; immunosuppressive; pericarditis;  
KM antiparkinsonian; ophthalmological; cell proliferative disorder;  
KM arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; angina pectoris;  
KM autoimmune disorder; AIDS; Crohn's disease; multiple sclerosis; epilepsy;  
KM ulcerative colitis; cardiovascular disorder; myocardial infarction;  
KM Raynaud's disease; myocarditis; neurological disorder; cataract;  
KM Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;  
KM developmental disorder; Duchenne muscular dystrophy; antipsoriatic;  
KM Becker muscular dystrophy; Cushing's syndrome.  
XX

OS Homo sapiens.  
XX  
XX WO200279441-A2.  
XX  
XX 10-OCT-2002.  
XX  
XX 29-MAR-2002; 2002MO-US009820.  
XX  
XX 30-MAR-2001; 2001US-0280527P.  
XX 06-APR-2001; 2001US-0282112P.  
XX 09-APR-2001; 2001US-0282702P.  
XX 13-APR-2001; 2001US-0283855P.  
XX 19-OCT-2001; 2001US-0343718P.  
XX 07-DEC-2001; 2001US-0339236P.  
XX 13-FEB-2002; 2002US-0357002P.  
XX  
XX (INCYTE) INCYTE GENOMICS INC.  
XX  
XX Baughn MR, Burford N, Ding L, Duggan BM, Elliott VS, Forsythe IU;  
XX Gandhi AR, Gietzen KU, Griffin JA, He A, Honchall CD, Ison CH;  
XX Lal PG, Lee EA, Lee S, Lu DM, Mason PM, Sanjanwala MM;  
XX Swarnakar A, Ramkumar J, Tang YT, Thangavelu K, Tran UK, Walla NK;  
XX Warren BA, Yao MG, Xu Y, Yue H;  
XX WPI; 2003-058429/05.  
XX N-PSDB; ABS57565.  
XX  
XX Novel human secreted protein useful for treating, preventing or  
PT diagnosing cancer, hepatitis, psoriasis, asthma, diabetes mellitus,  
PT anemia, epilepsy, cataract, Alzheimer's disease.  
XX  
XX Claim 76; Page 167; 188pp; English.  
XX  
XX This invention describes novel secreted proteins (SECP) which have  
CC antiarteriosclerotic, antiatherosclerotic, hepatotropic, cytosarctic, anti  
CC -HIV, antiallergic, antiasthmatic, antianaemic, antidiabetic,  
CC antiinflammatory, neuroprotective, anticancer, antipsoriatic, vasotropic,  
CC antineumatic, antiarthritic, cardiatic, hypotensive, anticonvulsant,  
CC nootropic, immunosuppressive, antiparkinsonian and ophthalmological  
CC activity. The polynucleotides and polypeptides of the invention can be  
CC used for diagnosing, treating or preventing cell proliferative disorder  
CC e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, cancer,  
CC autoimmune/inflammatory disorders e.g. acquired immunodeficiency syndrome  
CC (AIDS), allergies, asthma, anaemia, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, ulcerative colitis, psoriasis, rheumatoid arthritis,  
CC etc; cardiovascular disorder e.g. myocardial infarction, angina pectoris,  
CC hypertension, Raynaud's disease, myocarditis, pericarditis, etc;  
CC neurological disorders e.g. epilepsy, Huntington's disease, Parkinson's  
CC disease, Alzheimer's disease, Creutzfeldt-Jakob disease, etc; and  
CC developmental disorders e.g. Duchenne and Becker muscular dystrophy,  
CC cataract, gonadal dysgenesis, Cushing's syndrome, etc. The products of  
CC the invention can also be used for drug screening, proteome analysis,  
CC microarrays creating knock-in humanised animals or transgenic animals to  
CC model human diseases, in somatic or germline gene therapy, to generate a  
CC transcript image of a tissue or cell type, for detecting differences in  
CC the chromosomal location due to translocation, inversion, etc., among  
CC normal, carrier or affected individuals, and as hybridization probes for  
CC mapping naturally occurring genomic sequences. ABB84649-ABB84673  
CC represent secreted proteins encoded by the CDNA's shown in ABS57545-  
CC ABS57569, described in the disclosure of the invention  
XX  
XX Sequence 132 AA;  
SQ  
Query Match 17.5%; Score 110.5; DB 6; Length 132;  
Best Local Similarity 31.9%; Pred. No. 0.00042;  
Matches 30; Conservative 17; Mismatches 34; Indels 13; Gaps 3;  
QY 26 RSIPRGS-----FNRGAGAPGPGPPROQAGARLGAQSEFNDLRLQVNMGPQWHL 81  
DB 42 RQRPKESQDVDDSDNHTGEPVGDYK-----KMGTL---FGRLNKLINMGFTWTF 92  
QY 82 GNRHAYEPYTSILLFLMMLGVRLILVGLVYL 115

DB 93 GERIVEPVIVFFWMLMFLGLQALGLVAVLCV 126

## RESULT 13

ID ADE08064 standard; protein; 132 AA.

XX ADE08064;

XX 29-JAN-2004 (first entry)

XX Novel protein (useful for identifying genetic disorders) #219.

XX novel gene; novel protein; tissue marker; molecular weight marker;

XX chromosome marker; genetic disorder.

XX Unidentified.

XX WO2003054152-A2.

XX 03-JUL-2003.

XX 10-DEC-2002; 2002WO-US039555.

XX 10-DEC-2001; 2001US-0339739P.

XX 11-DEC-2001; 2001US-0339453P.

XX 14-MAR-2002; 2002US-0365091P.

XX 12-APR-2002; 2002US-0372381P.

XX 12-APR-2002; 2002US-0372615P.

XX 22-APR-2002; 2002US-00128558.

XX 24-APR-2002; 2002US-0376045P.

XX (HYSE-) HYSEQ INC.

XX Tang YF, Asundi V, Goodrich RM, Ren P, Zhang J, Zhao QA, Wang J;

XX Ghosh M, Xue AJ, Wehman T, Meng G, Zhou P, Drmanac RT, Wang Z;

XX Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

XX WPI; 2003-569235/53.

XX N-PSDB; ADE07153.

XX New polynucleotides, useful for expressing recombinant proteins for

XX analysis, characterization or therapeutic use, or as markers for tissues

XX in which the corresponding protein is preferentially expressed.

XX Claim 20; SEQ ID NO 1130; 1177BP; English.

XX The invention comprises the amino acid and coding sequences of novel

XX proteins. The DNA and protein sequences of the invention are useful as:

XX markers for tissues in which the corresponding protein is preferentially

XX expressed; as molecular weight markers on gels; as chromosome markers or

XX tags; to identify chromosomes or to map related gene positions; and to

XX compare with endogenous DNA sequences in patients to identify potential

XX genetic disorders. The present amino acid sequence represents a protein

XX of the invention.

XX Sequence 132 AA;

XX Query Match 17.5%; Score 110.5; DB 7; Length 132;

XX Best Local Similarity 31.9%; Pred. No. 0.00042;

XX Matches 30; Conservative 17; Mismatches 34; Indels 13; Gaps 3;

XX 26 RGSIRROG-----FRFGHGAPGCGPPRQQAARLGAAGSPFNDLRQLVNMGFPOWHL 81

XX 42 RGQRKESQDVESQNHTEGVEVDYK-----KMGTL---FSELNKLINMGFTMYF 92

XX 82 GNHAYEPVYSILLFLMMGLVGRGLLVGLVLYV 115

XX 93 GERIVEPVIVFFWMLMFLGLQALGLVAVLCV 126

AA677595  
ID AA677595 standard; protein; 92 AA.

XX AA677595;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:8361.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX colorectal carcinoma.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.

XX 29-SEP-1999; 99US-0157137P.

XX 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AAH37002.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

XX useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 11; Page 9643-9644; 9803BP; English.

XX AAH32943 to AAH37195 and AA677514 to AA677788 represent human colon

XX cancer-associated nucleic acid molecules (N) and proteins (P), where the

XX proteins are collectively known as colon cancer antigens. The colon

XX cancer antigens have cytosolic activity and can be used in gene therapy

XX and vaccine production. N and P may be used in the prevention, diagnosis

XX and treatment of diseases associated with inappropriate P expression. For

XX example, N and P may be used to treat disorders associated with decreased

XX expression by rectifying mutations or deletions in a patient's genome

XX that affect the activity of P by expressing inactive proteins or to

XX supplement the patient's own production of P. Additionally, N may be used

XX to produce the colon cancer-associated P, by inserting the nucleic acids

XX into a host cell and culturing the cell to express the proteins. N and P

XX can be used in the prevention, diagnosis and treatment of colorectal

XX carcinomas and cancers. AAH37196 to AAH37204 and AA677789 represent

XX sequences used in the expression/inactivation of the present invention. N.B.

XX Pages 666 to 682 and page 7053 of the sequence listing were missing at

XX time of publication, meaning no sequences are present for SEQ ID NO:1027

XX to 1052, 7921 and 7922

XX Sequence 92 AA;

XX Query Match 12.7%; Score 80; DB 4; Length 92;

XX Best Local Similarity 35.7%; Pred. No. 0.64;

XX Matches 30; Conservative 2; Mismatches 26; Indels 26; Gaps 4;

XX 15 DRRVR-----TTTPRGSIPIRQGFNRHGAPGCGPPRQQAARLGAAGSPFND 66

XX 15 DRRVRSSSGSPVLYRVRAS-PAPEGFASGHGALPGAGLR---AGHERGAAR----- 66

XX 67 LNRQLVNMGFPOWHLGNHAYEPVT 90

XX 67 -----PRAHGGRSAARPGT 80

XX RESULT 15

XX AAB42404

XX ID AAB42404 standard; protein; 325 AA.



AA042404;  
08-FEB-2001 (first entry)  
Human ORFX ORF2168 polypeptide sequence SEQ ID NO:4336.  
Human, open reading frame; ORFX; detection; cytostatic; hepatotropic;  
vulnary; antipsoptic; antiparkinsonian; nootropic; neuroprotective;  
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
immunostimulant; thrombolytic; coagulant; vasotrophic; antidiabetic;  
hypotensive; dermatological; immunosuppressive; antihypertensive;  
antiviral; antibacterial; antifungal; antineumatic; antihypertensive;  
antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
neurodegenerative disorder; osteoarthritis; graft vs host disease;  
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
cholesterol ester storage; systemic lupus erythematosus; infection;  
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
bone damage; cartilage damage; antineoplastic disease; coagulation;  
thrombosis; contraceptive.  
Homo sapiens.  
WO200058473-A2.  
05-OCT-2000.  
31-MAR-2000; 2000MO-US008621.  
31-MAR-1999; 99US-0127607P.  
02-APR-1999; 99US-0127636P.  
05-APR-1999; 99US-0127728P.  
30-MAR-2000; 2000US-00540763.  
(CURA-) CURAGEN CORP.  
Shinkets RA, Leach M;  
WPI: 2000-602362/57.  
N-PSDB; AAC76613.  
Novel nucleic acids and peptides derived from open reading frame X,  
useful for treating e.g. cancers, proliferative disorders,  
neurodegenerative disorders and cardiovascular disease.  
Claim 11; Page 3531-3532; 5507pp; English.  
AAC74446 to AAC77606 encode the proteins given in AA04237 to AA043397,  
which represent the human ORFX open reading frames 1 to 3161. The ORFX  
sequences have activities such as: cytostatic; hepatotropic; vulnary;  
antipsoptic; antiparkinsonian; nootropic; neuroprotective; osteopathic;  
anticonvulsant; antiparkinsonian; immunosuppressant; immunostimulant;  
cardiant; thrombolytic; coagulant; vasotrophic; antidiabetic; hypotensive;  
dermatological; immunosuppressive; antihypertensive; antineumatic;  
antiviral; antifungal; antineumatic; antihypertensive; antianemic. The  
sequences can be used for determining the presence of or predisposition  
to, or preventing or treating pathological conditions associated with an  
ORFX-associated disorder. The nucleic acids can be used to express ORFX  
proteins in gene therapy vectors. The proteins and nucleic acids may be  
used to treat cancers, proliferative disorders, neurodegenerative  
disorders, osteoarthritis, graft vs host disease, cardiovascular disease,  
diabetes mellitus, hypertension, hypothyroidism, cholesterol ester  
storage, systemic lupus erythematosus, severe combined immunodeficiency  
(SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune  
disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and  
cartilage damage, nocturnal haemoglobinuria, antineoplastic disease; to  
enhance coagulation; to inhibit thrombosis; and as a contraceptive

Sequence 325 AA;

Query Match 12.7%; Score 80; DB 3; Length 325;  
Best Local Similarity 35.5%; Pred. No. 2.8;  
Matches 27; Conservative 13; Mismatches 30; Indels 6; Gaps 3;

Qy 48 PPOQAGARLGAAGSPFNDLNRDLVNM--GFPQWHLGNHVEPVTSLILFLMLGVG 105  
Db 28 PFQALGALVNVQIFL--LLGQLVAVVAKYTRDHVGSFMT--SONLSTHLLILYVG 83  
Qy 106 ILLVGLVYLVSHTLSOR 121  
Db 84 LITFGLVYLVSHTLSOR 99

Search completed: August 24, 2004, 09:18:12  
Job time : 61 secs

**This Page Blank (uspto)**

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:17:05 ; Search time 15 seconds  
(without alignments)

775,945 Million cell updates/sec

Title: US-10-001-885-125

Sequence: 1 MVRILANGRIYQDDPRVRT.....GVKGLLVGLVYVSHLSQR 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*

1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83.5	13.2	219	2 T46472	hypothetical prote
2	81	12.8	475	2 A87396	sensor histidine k
3	77.5	12.3	692	2 T21095	hypothetical prote
4	76.5	12.1	1201	2 A83007	hypothetical prote
5	76	12.0	159	2 UQ0145	hypothetical 17.3K
6	74.5	11.8	1151	2 T18297	zinc-finger protei
7	73	11.6	205	1 SAVIMC	delta large antige
8	73	11.6	651	2 T46050	hypothetical prote
9	73	11.6	841	2 T01046	hypothetical prote
10	72.5	11.5	195	1 SAVILI	delta large antige
11	72.5	11.5	207	2 UQ2167	delta large antige
12	72.5	11.5	433	2 T36122	probable integral
13	72	11.4	214	1 SAVIDV	delta large antige
14	71.5	11.3	552	1 CREBPT	aspartate chemorec
15	71.5	11.3	553	2 AC0746	methyl-accepting c
16	71	11.2	345	2 A40616	34k antigen - Myco
17	71	11.2	345	2 A53138	gas1 homolog - hum
18	71	11.2	618	2 T42664	hypothetical prote
19	70.5	11.2	197	2 S22381	hypothetical prote
20	70	11.1	146	2 E70557	hypothetical prote
21	70	11.1	398	2 T34947	hypothetical prote
22	70	11.1	2559	2 T09144	probable guanine n
23	69.5	11.0	231	2 S75387	hypothetical prote
24	69.5	11.0	269	2 T42919	protein t10 - atel
25	69.5	11.0	447	2 T50705	gamma-aminobutyrat
26	69.5	11.0	530	2 F87590	serine proteinase
27	69	10.9	214	1 SAVILN	delta large antige
28	69	10.9	255	2 B87585	transcription regu
29	69	10.9	283	2 G01926	insulin promoter I

30	69	10.9	484	2 AG2097	serine/threonine k
31	69	10.9	493	2 F75113	NADH dehydrogenase
32	69	10.9	495	2 G71019	hypothetical prote
33	69	10.9	1454	2 B84535	probable retroelem
34	68.5	10.8	148	2 S74589	hypothetical prote
35	68.5	10.8	195	2 SAVIMC	delta large antige
36	68.5	10.8	195	2 S53112	delta antigen - he
37	68.5	10.8	206	2 T37094	hypothetical prote
38	68.5	10.8	454	1 C45340	nucleosacpid prote
39	68.5	10.8	653	2 T34356	hypothetical prote
40	68	10.8	530	2 T48627	hypothetical prote
41	67.5	10.7	525	2 T44445	chitinase (EC 3.2.
42	67.5	10.7	1747	2 A54121	collagen alpha-4 c
43	67	10.6	260	2 S22373	proline-rich prote
44	67	10.6	914	2 T52220	hypothetical prote
45	67	10.6	1121	2 A82809	exodeoxyribonuclea

#### ALIGNMENTS

RESULT 1  
T46472  
hypothetical protein DKFP434N2030.1 - human (fragment)  
C/Species: Homo sapiens (man)  
C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C/Accession: T46472  
R/Blum, H.; Bauersachs, S.; Nemes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A/Reference number: Z33034  
A/Accession: T46472  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-219 <AAA>  
A/Cross-references: EMBL:AL137488  
A/Experimental source: adult testis; clone DKFP434N2030  
C/Genetics:  
A/Note: DKFP434N2030.1

Query Match 13.2%; Score 83.5; DB 2; Length 219;  
Best Local Similarity 36.5%; Pred. No. 0.56;  
Matches 23; Conservativity 11; Mismatches 24; Indels 5; Gaps 2;

QY 41 APPGPGPRCOAGARLGAOSPFNDLNRQLVNMGFPQMHIGNHAYEVTSLILFLMM 100  
DB 9 APPRGAGSVRRRG-RLGALRLPC---RPGRSVGFPEGVNMGSSVRELVLFVLR 63  
QY 101 LGV 103  
DB 64 EGI 66

RESULT 2  
A87396  
sensor histidine kinase [imported] - Caulobacter crescentus  
C/Species: Caulobacter crescentus  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C/Accession: A87396  
R/Heiman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Neilson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Taub, R.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A/Title: Complete Genome Sequence of Caulobacter crescentus.  
A/Reference number: A87249; MUID:21173698; PMID:11259647  
A/Accession: A87396  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-475 <STO>  
A/Cross-references: GB:AE005673; NID:g13422503; PIDN:AAK23165.1; GSPDB:GN00148  
C/Genetics:  
A/Note: CC1181

Query Match 12.8%; Score 81; DB 2; Length 475;



RESULT 7  
 SAVLMC  
 delta large antigen - hepatitis delta virus (woodchuck isolate)  
 N:Alternate names: HDag  
 C:Species: hepatitis delta virus  
 C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 01-Aug-1997  
 C/Accession: A30054  
 R:Kuo, M.Y.P.; Goldberg, J.; Coates, L.; Mason, W.; Gerin, J.; Taylor, J.  
 J. Virol. 62, 1955-1961, 1988  
 A:Title: Molecular cloning of hepatitis delta virus RNA from an infected woodchuck liver  
 A:Reference number: A30054; MUID:88215010; PMID:3367426  
 A:Accession: A30054  
 A:Molecule type: genomic RNA  
 A:Residues: 1-205 <KIO>  
 A/Cross-references: GB:M21012  
 C/Comment: This virus is a replication-defective hepatitis B virus.  
 C:Superfamily: hepatitis delta virus large antigen  
 C:Keywords: core protein  
 F/2-205/Product: delta large antigen #status predicted <MAT>

Query Match 11.6%; Score 73; DB 1; Length 205;  
 Best Local Similarity 30.7%; Pred. No. 5.8;  
 Matches 23; Conservative 8; Mismatches 32; Indels 12; Gaps 4;

QY 10 IVODDDPRVATTTPPGSGIPRQSFNRGHPGPGPGPQQAGARLGAASPFNDLNR 69  
 DB 133 LTEDERRRRRVAGPVGCV--NPLEGSRGAPGGGFFVNLQ-----GVPSPPSRIGE 184

QY 70 QI---VNMGPQVNL 81  
 DB 185 GLDIRNGGFP-WDI 198

RESULT 8  
 T46050  
 hypothetical protein T16K5.190 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C/Accession: T46050  
 R:Reger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke, submitted to the protein sequence database, January 2000  
 A:Reference number: Z23015  
 A:Accession: T46050  
 A:Molecule type: DNA  
 A:Status: preliminary  
 A:Residues: 1-651 <RTE>  
 A/Cross-references: EMBL:A1132965  
 A:Experimental source: cultivar Columbia; BAC clone T16K5  
 C:Genetics:  
 A:Map position: 3  
 A:Insertions: 52/1; 83/1; 153/2; 209/1; 257/3; 278/3; 309/3; 363/3; 424/1; 479/1; 501/1; 54  
 A:Note: T16K5.190

Query Match 11.6%; Score 73; DB 2; Length 651;  
 Best Local Similarity 27.0%; Pred. No. 18;  
 Matches 17; Conservative 10; Mismatches 36; Indels 0; Gaps 0;

QY 15 DPRVATTTPPGSGIPRQSFNRGHPGPGPGPQQAGARLGAASPFNDLNRQLVNM 74  
 DB 532 DPGHPVAPPPGCGPRKEGYPAGYPPAGYPPQYQAGYPPAGYPPQGGGCGVPA 591

QY 75 GFP 77  
 DB 592 GYP 594

RESULT 9  
 T01046  
 hypothetical protein YUP8H12R.29 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 22-Oct-1999  
 C/Accession: T01046  
 R:Theologis, A.; Vysotskaia, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; Kwan

Oefner, P.; Davis, R.W.  
 submitted to the EMBL Data Library, May 1998  
 A:Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.  
 A:Reference number: Z14227  
 A:Accession: T01046  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-841 <THB>  
 A/Cross-references: EMBL:AC002986; NID:g2494106; PID:g3152568; GSPDB:GN00059; ATSP:YUP8H  
 C:Genetics:  
 A:Gene: ATSP:YUP8H12R.29  
 A:Map position: 1  
 A:Insertions: 74/3; 93/3; 164/1

Query Match 11.6%; Score 73; DB 2; Length 841;  
 Best Local Similarity 35.8%; Pred. No. 24;  
 Matches 24; Conservative 8; Mismatches 19; Indels 16; Gaps 5;

QY 29 IPQSFNRGHPGPGPGPQQAGARLGAASPFNDLNRQLVNM--GF 76  
 DB 241 VPKSPFVSX---PPPGSIPQRLGHPNIPYQSGPQWGSFNPFPNLQPLPSMHGS 297

QY 77 PGMHGN 83  
 DB 298 PQ-HTGN 303

RESULT 10  
 SAVLMC  
 delta large antigen - hepatitis delta virus (strain Lebanon-1)  
 N:Alternate names: HDag  
 C:Species: hepatitis delta virus  
 A:Note: host Homo sapiens (man)  
 C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 01-Aug-1997  
 C/Accession: A40247  
 R:Lee, C.M.; Bih, F.Y.; Chao, Y.C.; Govindarajan, S.; Lai, M.M.C.  
 Virology 188, 265-273, 1992  
 A:Title: Evolution of hepatitis delta virus RNA during chronic infection.  
 A:Reference number: A40247; MUID:92230225; PMID:1566577  
 A:Accession: A40247  
 A:Molecule type: genomic RNA  
 A:Residues: 1-195 <LBB>  
 A/Cross-references: GB:M64917  
 C/Comment: This virus is a replication-defective hepatitis B virus.  
 C:Superfamily: hepatitis delta virus large antigen  
 C:Keywords: core protein  
 F/2-195/Product: delta large antigen #status predicted <MAT>

Query Match 11.5%; Score 72.5; DB 1; Length 195;  
 Best Local Similarity 31.5%; Pred. No. 6.1;  
 Matches 23; Conservative 8; Mismatches 31; Indels 11; Gaps 3;

QY 8 GRIVODDDPRVATTTPPGSGIPRQSFNRGHPGPGPGPQQAGARLGAASPFNDL 67  
 DB 131 GLTTEDEKRRRVAGPPTGVNPLEGQR--GAPGGGFFVSMQ-----GVPSPPFRRH 182

QY 68 NFQL---VNMGP 77  
 DB 183 GEGLDARDGRFP 195

RESULT 11  
 JQ2167  
 delta large antigen - hepatitis delta virus (strain CAR)  
 N:Alternate names: HDag  
 C:Species: hepatitis delta virus  
 C>Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 07-May-1999  
 C/Accession: JQ2167  
 R:Tang, J.R.; Hantz, O.; Vitvileki, L.; Lamelin, J.P.; Parana, R.; Cova, L.; Lesbordes, U.; Gen. Virol. 74, 1827-1835, 1993  
 A:Title: Discovery of a novel point mutation changing the HDag expression of a hepatitis A:Reference number: JQ2167; MUID:93389437; PMID:8376962  
 A:Accession: JQ2167

A:Molecule type: mRNA  
A:Residues: 1-207 <TAN>  
A:Experimental source: isolate CAR  
A>Note: the authors translated the codon CGA for residue 85 as nothing and AGA for residue 86 as nothing.  
C:Comment: This protein is one of hepatitis delta virus specific-proteins.  
C:Superfamily: hepatitis delta virus large antigen  
C:Keywords: core protein, lipoprotein, methylated carboxyl end; prenylated cysteine  
F/204/Binding site: farnesyl (Cys) (covalent) #status predicted  
F/204/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

Query Match 11.5%; Score 72.5; DB 2; Length 207;  
Best Local Similarity 30.6%; Pred. No. 6.5; Mismatches 31; Indels 11; Gaps 3;  
Matches 22; Conservative 8;

QY 10 IVQDDPRVTRTTTPPGSGIPRGSFNRGCGAPGPGPPRQQAAGARLGAAGSPFNDLNR 69  
DB 126 LTEDERRERRRRAAGPVGAV--NPLEGSGRAGPGGFPVNLQ-----GVPSFPFRTGE 177  
QY 70 QL---VVMGFPFQ 78  
DB 178 GLDIRGNQGFPEK 189

## RESULT 12

T36122

Probable integral membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C&gt;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C/Accession: T36122

R/Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1999

A:Reference number: Z21597

A/Accession: T36122

A/Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-433 &lt;MOR&gt;

A:Cross-references: EMBL:AL049707; PIDN:CA841288.1; GSPDB:GN00070; SCODEB:SCB15.19

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCODEB:SCB15.19

Query Match 11.5%; Score 72.5; DB 2; Length 433;  
Best Local Similarity 31.3%; Pred. No. 14;

Matches 26; Conservative 9; Mismatches 31; Indels 17; Gaps 5;

QY 11 VQDDPRVTRTTTPPGSGIPRGSFNRGCGAPGPGPPRQQAAGARLGAAGSPFNDLNR 68  
DB 241 LQETDPQAPAGEQ--RAELVKGPEADRPAPPSGPGSGPQAGGEDPQSPWPN--T 296

QY 69 RQLVMMG-----FPQWH 80  
DB 297 RQVAYQGYDPSYDPGQVQWY 319

## RESULT 13

SAV1DV

delta large antigen - hepatitis delta virus (chimpanzee isolate)

N/Alternate names: HDAG

C:Species: hepatitis delta virus

C&gt;Date: 30-Sep-1987 #sequence\_revision 31-Dec-1990 #text\_change 01-Aug-1997

C/Accession: A26176; A30028

R/Wang, K.S.; Choo, Q.L.; Weiner, A.J.; Ou, J.H.; Najarian, R.C.; Thayer, R.M.; Mullerbach

Nature 323, 508-514, 1986

A&gt;Title: Structure, sequence and expression of the hepatitis delta (delta) viral genome.

A:Reference number: A26176; MUID:87014821; PMID:3762705

A/Accession: A26176

A:Molecule type: genomic RNA

A:Residues: 1-161 &lt;MAN&gt;

A/Cross-references: GB:X04451

A&gt;Note: the nucleotide sequence contains a frameshift error in codon 162 that has been corrected.

R/Wang, K.S.; Choo, Q.L.; Weiner, A.J.; Ou, J.H.; Najarian, R.C.; Thayer, R.M.; Mullerbach

Nature 328, 456, 1987

Reference number: X00078

A:Contents: corrections

A:Accession: A30028

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: genomic RNA

A:Residues: 1-214 &lt;MA2&gt;

A/Cross-references: GB:M21012

A&gt;Note: this corrects a frameshift error in codon 162 of the published sequence

C:Comment: This virus is a replication-defective hepatitis B virus.

C:Superfamily: hepatitis delta virus large antigen

C:Keywords: core protein, lipoprotein, methylated carboxyl end; prenylated cysteine

F/2-211/Product: delta large antigen #status predicted &lt;MA2&gt;

F/211/Binding site: farnesyl (Cys) (covalent) #status predicted

F/211/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

Query Match 11.4%; Score 72; DB 1; Length 214;  
Best Local Similarity 32.9%; Pred. No. 7.6;

Matches 25; Conservative 8; Mismatches 29; Indels 14; Gaps 5;

QY 10 IVQDDPRVTRTTTPPGSGIPRGSFNRGCGAPGPGPPRQQAAGARLGAAGSPFNDLNR 68  
DB 133 LTEDERRERRRRAAGPVGAVPLR---GSGRAGGGGFPVSLQ-----GVPSFPFRTGE 183

QY 69 RQL---VVMGFPFQ 81  
DB 184 EGLDIRGNRGP-NDI 198

## RESULT 14

Q8BBDT

aspartate chemoreceptor protein - Salmonella typhimurium

N/Alternate names: methyl-accepting chemotaxis protein II (MCP-II)

C:Species: Salmonella typhimurium

C&gt;Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 15-Aug-1997

C/Accession: A03441; A29053

R/Russo, A.F.; Koshland Jr., D.E.

Science 220, 1016-1020, 1983

A&gt;Title: Separation of signal transduction and adaptation functions of the aspartate receptor

A:Reference number: A03441; MUID:83197387; PMID:6302843

A/Accession: A03441

A:Molecule type: DNA

A:Residues: 1-552 &lt;RUS&gt;

A&gt;Note: residues 294-296, 301-302, 308-309, 483-484, 490-491, and 499-501 are potential N

R/Mowbray, S.L.; Foster, D.L.; Koshland Jr., D.E.

J. Biol. Chem. 260, 11711-11718, 1985

A&gt;Title: Chemolytic fragments identified with domains of the aspartate chemoreceptor.

A:Reference number: A29053; MUID:86008217; PMID:2955347

A/Accession: A29053

A:Molecule type: protein

A:Residues: 1-245, 'E', 247-552 &lt;MOW&gt;

C:Comment: This protein responds to changes in aspartate concentration in the environment

C:Genetics: various levels of methylation.

A:Gene: tar

C:Superfamily: methyl-accepting chemotaxis protein

C:Keywords: chemotaxis; methylated amino acid; sensory transduction; transmembrane protein

F/295,309/Modified site: glutamate methyl ester (Glu) (by cheB-dependent deamidation and

F/302,491,500/Modified site: glutamate methyl ester (Glu) #status predicted

Query Match 11.3%; Score 71.5; DB 1; Length 552;  
Best Local Similarity 28.2%; Pred. No. 22;

Matches 22; Conservative 15; Mismatches 28; Indels 13; Gaps 3;

QY 50 QQAGARLGAAGSPFNDLNLQVMMG-----FPQWHGNAVPPVSIILFLPMML--G 102  
DB 158 QNALGKALGNARVSEULVQTPQSGADYRFAQMDG-----VLAVALVILVMVWF 211  
QY 103 VQGLLVGLVYVSHSQ 120  
DB 212 IRRALNPLARVTHIRE 229

## RESULT 15

X00078



**This Page Blank (uspio)**



OM protein - protein search, using sw model

Run on: August 24, 2004, 09:17:04 (Search time 9 Seconds)

(Without alignments)  
700.054 Million cell updates/sec

Title: US-10-001-885-125

Perfect score: 632  
Sequence: 1 MVRILNGRIVQDDPRVRL.....GVRGLLVGLVYLVSLSQR 121

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	11.9	735	1	ABC8_HUMAN
2	73	11.6	735	1	ABC8_HUMAN
3	72.5	11.5	195	1	AAAT_HUMAN
4	72	11.4	195	1	AAAT_HUMAN
5	72	11.4	768	1	AD11_HUMAN
6	71.5	11.3	553	1	FA10_RABIT
7	71.5	11.3	553	1	MCPE2_SALTY
8	71	11.2	345	1	GAS1_HUMAN
9	71	11.2	478	1	CTP8_HUMAN
10	71	11.2	481	1	CTP8_HUMAN
11	70.5	11.1	1054	1	S24A_ARATH
12	70	11.1	214	1	AAAT_HUMAN
13	69	10.9	283	1	IPF1_HUMAN
14	69	10.9	283	1	IPF1_HUMAN
15	69	10.8	195	1	AAAT_HUMAN
16	68.5	10.8	195	1	AAAT_HUMAN
17	68.5	10.8	195	1	AAAT_HUMAN
18	68.5	10.8	195	1	AAAT_HUMAN
19	68.5	10.8	195	1	AAAT_HUMAN
20	68	10.8	195	1	AAAT_HUMAN
21	67	10.6	2191	1	POLG_HUMAN
22	66.5	10.5	255	1	LYOX_HUMAN
23	66.5	10.5	420	1	LYOX_HUMAN
24	66.5	10.5	424	1	LYOX_HUMAN
25	66.5	10.5	738	1	AB10_HUMAN
26	66.5	10.5	738	1	AB10_HUMAN
27	66.5	10.5	1125	1	IF2_HUMAN
28	66	10.4	283	1	IPF1_RAT
29	66	10.4	628	1	TES1_RAT
30	66	10.4	705	1	HS70_HUMAN
31	66	10.4	1505	1	CUT1_HUMAN
32	65.5	10.4	286	1	PMP3_MOUSE
33	65.5	10.4	407	1	AMEN_MOUSE

34	65.5	10.4	1462	1	NC02_MOUSE
35	65	10.3	362	1	HRCA_BRAJA
36	65	10.3	385	1	ITR5_BOVIN
37	65	10.3	461	1	SYT_SALTY
38	65	10.3	461	1	SYT_SALTY
39	65	10.3	606	1	AAAF_PIG
40	65	10.3	666	1	ENV_MLVHO
41	65	10.3	674	1	CA1A_BOVIN
42	65	10.3	1596	1	MAV_DROME
43	64.5	10.2	297	1	PCW3_BROWH
44	64.5	10.2	361	1	SSB3_RAT
45	64.5	10.2	388	1	SSB3_HUMAN

#### ALIGNMENTS

RESULT 1  
ABC8\_HUMAN STANDARD; PRT; 735 AA.  
ID ABC8\_HUMAN  
AC Q9NUT2; Q95787;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE ATP-binding cassette, sub-family B, member 8, mitochondrial precursor  
DE (Mitochondrial ATP-binding cassette 1) (M-ABC1).  
GN ABC8 OR MAB1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN (1)  
RP SEQUENCE FROM N.A. (ISOFORM SHORT), AND MUTAGENESIS.  
RX MEDLINE=99096930; PubMed=9878413;  
RA Hogue D.L., Liu L., Ling V.;  
RT Identification and characterization of a mammalian mitochondrial ATP-  
binding cassette membrane protein.;  
RL J. Mol. Biol. 285:379-389 (1999).  
RN (2)  
RP SEQUENCE FROM N.A. (ISOFORM LONG).  
RC TISSUE=Placenta;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Ichibashi T., Fujimori K.,  
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,  
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K.,  
RA Masuko Y., Kanehori K.;  
RT "NEO human cDNA sequencing project";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN (3)  
RP VARIANT ILE-152.  
RX MEDLINE=21686803; PubMed=11829140;  
RA Saito S., Iida A., Sekine A., Miura Y., Ogawa C., Kawauchi S.,  
RA Higuchi S., Nakamura Y.;  
RT "Three hundred twenty-six genetic variations in genes encoding nine  
members of ATP-binding cassette, subfamily B (ABC8/MDR/TFP), in the  
Japanese population.";  
RL J. Hum. Genet. 47:38-50(2002).  
CC -!- SUPPNT: Monomer (Potential).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
inner membrane.  
CC -!- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Name=Long;  
isoId=Q9NUT2-1; Sequence=Displayed;  
Name=Short;  
isoId=Q9NUT2-2; Sequence=VSP\_000026;  
Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Ubiquitous.  
CC -!- SIMILARITY: Belongs to the ABC transporter family. MDR subfamily.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.lib-sib.ch/announce/>  
 CC or send an email to [license@lib-sib.ch](mailto:license@lib-sib.ch)).

-----  
 DR EMBL; AF047690; AAD15748.1; -  
 DR EMBL; AK002018; BAA92038.1; -  
 DR HSSP; P13569; INBD  
 DR Genew; HGNC:49; ABCB8.  
 DR MIM; 605464; -  
 DR GO; GO:0005624; C:membrane fraction; TAS.  
 DR GO; GO:0005740; C:mitochondrial membrane; TAS.  
 DR GO; GO:0005524; F:ATP binding; TAS.  
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; TAS.  
 DR GO; GO:0005215; F:transporter activity; TAS.  
 DR GO; GO:0006810; P:transport; TAS.  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR001140; ABC TM transp.  
 DR InterPro; IPR003439; ABC transporter.  
 DR Pfam; PF00664; ABC membrane; 1.  
 DR Pfam; PF00005; ABC tran; 1.  
 DR ProDom; PD000006; ABC transporter; 1.  
 DR SMART; SMO0382; AAA; 1.  
 DR PROSITE; PS050929; ABC\_TM1F; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
 DR PROSITE; PS00893; ABC\_TRANSPORTER\_2; 1.  
 KM ATP-binding; Transmembrane; Transport; Mitochondrion; Inner membrane;  
 KW Transist peptide; Alternative splicing; Polymorphism.  
 FT TRANSIT ? 735  
 FT CHAIN ? 735  
 FT TRANSMEM 86 102  
 FT TRANSMEM 145 165  
 FT TRANSMEM 196 216  
 FT TRANSMEM 296 316  
 FT TRANSMEM 383 403  
 FT NP BIND 507 514  
 FT VASPLIC 33 49  
 FT VARIANT 152 152  
 FT MUTAGEN 512 513  
 FT CONFLICT 174 174  
 FT CONFLICT 239 239  
 FT CONFLICT 278 278  
 FT CONFLICT 326 326  
 FT CONFLICT 347 347  
 FT CONFLICT 365 365  
 FT SEQUENCE 735 AA; 79947 MW; 879519858F7EC7A6 CRC64;

Query Match 11.9%; Score 75; DB 1; Length 735;  
 Best Local Similarity 35.7%; Pred. No. 9.9;  
 Matches 25; Conservative 13; Mismatches 26; Indels 6; Gaps 3;

Db 157 GAALVNVQITPL--LLGQLVKVAKYTRDVGSPMTT--SONLSTHLLILIGVGLLTFTG 212

QY 112 VYLVSHLSOR 121  
 : : : : :  
 : : : : :  
 Db 213 LVLLSHVGER 222

RESULT 2  
 AANT HDVMO STANDARD; PRT; 205 AA.  
 ID AANT HDVMO STANDARD; PRT; 205 AA.  
 AC P29997;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 CC Delta antigen.  
 CC Hepatitis delta virus (isolate Woodchuck) (HDV).

OX NCBI\_TaxID=31764;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86215010; PubMed=3367426;  
 RA Kuo W.Y.F., Goldberg J., Coates U., Mason W., Gerin J., Taylor J.,  
 RT "Molecular cloning of hepatitis delta virus RNA from an infected  
 RT woodchuck liver: sequence, structure, and applications.";  
 RL J. Virol. 62:1855-1861(1988).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- PTM: Phosphorylated.  
 CC -1- DISEASE: BINDS SPECIFICALLY TO ANTISERA FROM PATIENTS WITH CHRONIC  
 CC HEPATITIS DELTA VIRAL INFECTIONS.  
 CC -1- MISCELLANEOUS: Delta antigen binds specifically to HDV RNA.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.lib-sib.ch/announce/>  
 CC or send an email to [license@lib-sib.ch](mailto:license@lib-sib.ch)).

-----  
 DR EMBL; M21012; AAA45723.1; ALT\_TERM.  
 DR PIR; A30054; SAVIMC.  
 DR InterPro; IPR002506; HDV\_ag.  
 DR Pfam; PF01517; HDV\_ag; 1.  
 DR ProDom; PD002887; HDV\_ag; 1.  
 KW Antigen; RNA-binding; Nuclear protein; Phosphorylation.  
 SQ SEQUENCE 205 AA; 23047 MW; 8B33C73B7872B98 CRC64;

Query Match 11.6%; Score 73; DB 1; Length 205;  
 Best Local Similarity 30.7%; Pred. No. 4.1;  
 Matches 23; Conservative 8; Mismatches 32; Indels 12; Gaps 4;

QY 10 IVODDDPRVETTTOPRGRSIPROSGFNRGAPGCGPGRQCGARLGAASPFNDLR 69  
 : : : : :  
 : : : : :  
 Db 133 LTBEERREKRVVAGPVGCV--NPLEGSRKAPGCGGVPMNQ-----GVPESEFSTGE 164

QY 70 QL--VNMGPFQWHL 81  
 : : : : :  
 : : : : :  
 Db 185 GLDIRNGQGP-WDI 198

RESULT 3  
 AANT HDVLI STANDARD; PRT; 195 AA.  
 ID AANT HDVLI STANDARD; PRT; 195 AA.  
 AC P29833;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 CC Delta antigen.  
 CC Hepatitis delta virus (isolate Lebanon-1) (HDV).  
 CC Viruses; Deltavirus.  
 OX NCBI\_TaxID=31763;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92230225; PubMed=1566577;  
 RA Lee C.M., Bin F.Y., Chao Y.C., Govindarajan S., Lai M.M.C.,  
 RT "Evolution of hepatitis delta virus RNA during chronic infection.";  
 RT Virology 188:265-273(1992).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- PTM: Phosphorylated.  
 CC -1- DISEASE: BINDS SPECIFICALLY TO ANTISERA FROM PATIENTS WITH CHRONIC  
 CC HEPATITIS DELTA VIRAL INFECTIONS.  
 CC -1- MISCELLANEOUS: Delta antigen binds specifically to HDV RNA.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.lib-sib.ch/announce/>

CC EMBL: M84917; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: A40247; SAVILL.  
 DR InterPro: IPR002506; HDV\_ag.  
 DR Pfam: PF01517; HDV\_ag; 1.  
 DR ProDom: PD002887; HDV\_ag; 1.  
 DR Antigen: RNA-binding; Nuclear protein; Phosphorylation.  
 SQ SEQUENCE 195 AA; 21846 MW; 288929C585B5D418 CRC64;  
 Query Match 11.5%; Score 72.5; DB 1; Length 195;  
 Best Local Similarity 31.5%; Pred. No. 4.3;  
 Matches 23; Conservative 8; Mismatches 31; Indels 11; Gaps 3;  
 QY 8 GEIVODDPRVTRTTTOPRGSIPROSFNKGHAPGPGPGPQQAGARUGAAGSPENDL 67  
 Db 131 GRLTEEDKRRKRRVAGPFTGGVNLGGGR--GAPGGGFPVPSNQ-----GVPSFPFRRH 182  
 QY 68 NRQL---VNMGPF 77  
 Db 183 GEGLDARGDRGFP 195  
 RESULT 4  
 AANT\_HDVIT STANDARD; PRT; 214 AA.  
 AC P06934;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE Delta antigen (Alpha antigen).  
 OS Hepatitis delta virus (isolate Italian) (HDV).  
 OC Viruses; Delta virus.  
 OX NCBI\_TaxID=10423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87014821; PubMed=3762705;  
 RA Wang K.S., Choo Q.L., Weiner A.J., Ou J.H., Najarian R.C.,  
 RA Thayer R.M., Mullendach G.T., Denniston K.J., Gerin J.L., Houghton M.;  
 RT "Structure, sequence and expression of the hepatitis delta (delta)  
 RT viral genome.";  
 RL Nature 323:508-514(1986).  
 RN [2]  
 RP REVISIONS TO 162-214.  
 RA Wang K.S., Choo Q.L., Weiner A.J., Ou J.H., Najarian R.C.,  
 RA Thayer R.M., Mullendach G.T., Denniston K.J., Gerin J.L., Houghton M.;  
 RL Nature 328:456-456(1987).  
 RN [3]  
 RP SUBCELLULAR LOCATION, AND PHOSPHORYLATION.  
 RX MEDLINE=88230601; PubMed=3373572;  
 RA Chang M.F., Baker S.C., Soe L.H., Kamahora T., Keck J.G., Makino S.,  
 RA Govindarajan S., Lai M.M.C.;  
 RT "Human hepatitis delta antigen is a nuclear phosphoprotein with RNA-  
 RT binding activity.";  
 RL J. Virol. 62:2403-2410(1988).  
 RN [4]  
 RP RNA-BINDING SPECIFICITY.  
 RX MEDLINE=90347794; PubMed=2200884;  
 RA Lin J.-H., Chang M.F., Baker S.C., Govindarajan S., Lai M.M.C.;  
 RT "Characterization of hepatitis delta antigen: specific binding to  
 RT hepatitis delta virus RNA.";  
 RL J. Virol. 64:4051-4058(1990).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- PTM: Phosphorylated.  
 CC -1- DISEASE: BINDS SPECIFICALLY TO ANTISERA FROM PATIENTS WITH CHRONIC  
 CC HEPATITIS DELTA VIRAL INFECTIONS.  
 CC -1- MISCELLANEOUS: Delta antigen binds specifically to HDV RNA.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC EMBL: X04451; CAA28050.1; ALT\_SEQ.  
 DR PIR: A26176; SAVILDV.  
 DR InterPro: IPR002506; HDV\_ag.  
 DR Pfam: PF01517; HDV\_ag; 1.  
 DR ProDom: PD002887; HDV\_ag; 1.  
 DR Antigen: RNA-binding; Nuclear protein; Phosphorylation.  
 SQ SEQUENCE 214 AA; 24078 MW; 83221BEEB0A2688 CRC64;  
 Query Match 11.4%; Score 72; DB 1; Length 214;  
 Best Local Similarity 32.9%; Pred. No. 5.3;  
 Matches 25; Conservative 8; Mismatches 29; Indels 14; Gaps 5;  
 QY 10 IVODDDPRVTRTTTOPRGSIPROSFNKGHAPGPGPGPQQAGARUGAAGSPENDL 68  
 Db 133 LREDEKRRKRRVAGPFTGGVNLGGGR--GAPGGGFPVPSLQ-----GVPSFPFRRH 183  
 QY 69 RQL---VNMGFPWH 81  
 Db 184 EGDIRGNRGFP-WDI 198  
 RESULT 5  
 AD11\_HUMAN STANDARD; PRT; 769 AA.  
 ID AD11\_HUMAN  
 AC Q75078; Q14808; Q14809; Q14810;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE ADAM 11 precursor (A disintegrin and metalloproteinase domain 11)  
 DE (Metalloproteinase-like, disintegrin-like, and cysteine-rich protein)  
 DE (MDC).  
 GN ADAM11 OR MDC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RX MEDLINE=98359734; PubMed=9693107;  
 RC TISSUE=Brain;  
 RA Sagane K., Ohya Y., Hasegawa Y., Tanaka I.;  
 RA "Metalloproteinase-like, disintegrin-like, cysteine-rich proteins MDC2  
 RT and MDC3: novel human cellular disintegrins highly expressed in the  
 RT brain.";  
 RL Biochem. J. 334:93-98(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
 RC TISSUE=Cerebellum;  
 RX MEDLINE=94073190; PubMed=8252040;  
 RA Eml M., Katagiri T., Harada Y., Saito H., Inazawa J., Ito I.,  
 RA Kasumi F., Nakamura Y.;  
 RT "A novel metalloproteinase/disintegrin-like gene at 17q21.3 is  
 RT somatically rearranged in two primary breast cancers.";  
 RL Nat. Genet. 5:151-157(1993).  
 RN [3]  
 RP SEQUENCE OF 106-769 FROM N.A. (ISOFORMS LONG AND SHORT).  
 RC TISSUE=Brain, Breast, Ovary, and Testis;  
 RX MEDLINE=95044425; PubMed=7956356;  
 RA Katagiri T., Harada Y., Eml M., Nakamura Y.;  
 RT "Human metalloproteinase/disintegrin-like (MDC) gene: exon-intron  
 RT organization and alternative splicing.";  
 RL Cytogenet. Cell Genet. 68:39-44(1995).  
 CC -1- FUNCTION: Probable ligand for integrin in the brain. This is a non  
 CC catalytic metalloproteinase-like protein.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Name=Long; Synonyms=MDC-769; Named isoforms=2;  
 CC Name=Short; Synonyms=MDC-524;  
 CC Name=Short; Synonyms=MDC-524;  
 CC IsoId=O75078-2; Sequence=VSP\_005472, VSP\_005473, VSP\_005474,

VSP\_005475;

-1- TISSUE SPECIFICITY: Expressed predominantly in brain. slightly detected or not at all in other tissues.

-1- DOMAIN: A conserved motif [AN(E/D)CD] within the disintegrin-like domain could be involved in the binding to the integrin receptor.

-1- PFM: The precursor is cleaved by a furin endopeptidase (By similarity).

-1- DISEASE: Rearrangements occur in breast and ovarian cancers which involve multiple exons and disrupt the coding region.

-1- SIMILARITY: Belongs to peptidase family M12B.

-1- SIMILARITY: Contains 1 EGF-like domain.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See [http://www.lbd-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@sib-sib.ch](http://www.lbd-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch)).

-----

CC CC EMBL; AB009675; BAA32352.1; -.  
CC CC EMBL; D17390; BAA04213.1; -.  
CC CC EMBL; D31872; BAA06670.1; -.  
CC CC EMBL; D31872; BAA06671.1; -.  
CC CC PIR; I65967; I65967.  
CC CC HSSP; P18619; IFTL.  
CC CC MEROPS; M12.976; -.  
CC CC Genew; HGNC:189; ADAM11.  
CC CC MIM; 155120; -.  
CC DR GO; GO:0005886; C:plasma membrane; TAS.  
CC DR GO; GO:0005178; F:integrin binding; TAS.  
CC DR GO; GO:0008237; F:metallopeptidase activity; TAS.  
CC DR GO; GO:0007229; P:integrin-mediated signaling pathway; TAS.  
CC DR InterPro; IPR006586; ADAM\_cysteine.  
CC DR InterPro; IPR001762; Disintegrin.  
CC DR InterPro; IPR006209; EGF like.  
CC DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
CC DR InterPro; IPR001590; Peptidase\_M12B.  
CC DR InterPro; IPR002870; Peptidase\_M12B\_N.  
CC DR Pfam; PF00200; disintegrin; 1.  
CC DR Pfam; PF01562; Peg\_M12B\_propep; 1.  
CC DR Pfam; PF01421; RegIolyein; 1.  
CC DR PRINTS; PR00289; DISINTEGRIN.  
CC DR ProDom; PD00664; Disintegrin; 1.  
CC DR SMART; SMO0608; ACRI.1.  
CC DR SMART; SMO0050; DISTR.1.  
CC DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
CC DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
CC DR PROSITE; PS50214; DISINTEGRIN\_2; 1.  
CC DR PROSITE; PS00022; EGF\_1; 1.  
CC DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
CC DR PROSITE; PS50026; EGF\_3; 1.  
CC DR PROSITE; PS00142; ZINC\_PROTEASE; FALSE\_NEG.  
CC KW Signal; Glycoprotein; Transmembrane; EGF-like domain;  
CC KW Alternate splicing.  
CC FT SIGNAL 1 23  
CC FT PROPEP 24 225  
CC FT CHAIN 226 769  
CC FT DOMAIN 226 734  
CC FT TRANSMEM 735 755  
CC FT DOMAIN 736 769  
CC FT DOMAIN 226 438  
CC FT DOMAIN 444 531  
CC FT DOMAIN 532 676  
CC FT DOMAIN 677 709  
CC FT DISULFID 349 433  
CC FT DISULFID 503 516  
CC FT DISULFID 677 692  
CC FT DISULFID 666 688  
CC FT DISULFID 700 709  
CC FT CAROYHD 96 96  
CC N-LINKED (GLCNAC... ) (POTENTIAL).  
CC BY SIMILARITY.  
CC EXTRACELLULAR (POTENTIAL).  
CC POTENTIAL.  
CC CYTOPLASMIC (POTENTIAL).  
CC METALLOPROTEASE-LIKE.  
CC DISINTEGRIN-LIKE.  
CC CYS-RICH.  
CC EGF-LIKE.  
CC BY SIMILARITY.  
CC POTENTIAL.  
CC BY SIMILARITY.  
CC BY SIMILARITY.  
CC BY SIMILARITY.  
CC BY SIMILARITY.  
CC N-LINKED (GLCNAC... ) (POTENTIAL).  
CC BY SIMILARITY.

[illegible]

[illegible]

RT bacterial aspartate receptor with and without a ligand.";  
 RL Science 254:1342-1347(1991).  
 RN (5)  
 RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 36-180.  
 RX MEDLINE=93252857; PubMed=8486661;  
 RA Yeh J.-I., Biemann H.-P., Pandit J., Koshland D.E., Kim S.-H.;  
 RT "The three-dimensional structure of the ligand-binding domain of a  
 RT wild-type bacterial chemotaxis receptor. Structural comparison to  
 RT the cross-linked mutant forms and conformational changes upon ligand  
 RT binding.";  
 RL J. Biol. Chem. 268:9787-9792(1993).  
 RN (6)  
 RN X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 26-180.  
 RX MEDLINE=96428686; PubMed=8831788;  
 RA Yeh J.-I., Biemann H.-P., Prive G.G., Pandit J., Koshland D.E. Jr.,  
 RA Kim S.-H.;  
 RT "High-resolution structures of the ligand binding domain of the wild-  
 RT type bacterial aspartate receptor.";  
 RL J. Mol. Biol. 262:186-201(1996).  
 CC -1- FUNCTION: RECEPTOR FOR THE ATTRACTANT L-ASPARTATE AND RELATED  
 CC AMINO AND DICARBOXYLIC ACIDS. TAR MEDIATES TAXIS AWAY FROM THE  
 CC REPULSANTS COBALT AND NICKEL. UNLIKE E. COLI TAR IT DOES NOT  
 CC MEDIATES MALTOSE TAXIS.  
 CC -1- FUNCTION: CHEMOTACTIC-SIGNAL TRANSDUCERS RESPOND TO CHANGES IN THE  
 CC CONCENTRATION OF ATTRACTANTS AND REPULSANTS IN THE ENVIRONMENT,  
 CC TRANSDUCE A SIGNAL FROM THE OUTSIDE TO THE INSIDE OF THE CELL, AND  
 CC FACILITATE SENSORY ADAPTATION THROUGH THE VARIATION OF THE LEVEL  
 CC OF METHYLATION. ATTRACTANTS INCREASE THE LEVEL OF METHYLATION  
 CC WHILE REPULSANTS DECREASE THE LEVEL OF METHYLATION. THE METHYL  
 CC GROUPS ARE ADDED BY THE METHYLTRANSFERASE CHEB. AND REMOVED BY THE  
 CC METHYLERASE CHEB.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
 CC -1- SIMILARITY: Contains 1 methyl-accepting transducer domain.  
 CC -1- SIMILARITY: Contains 1 HAMP domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J01809; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; AE008785; AAL20835.1; -.  
 DR PDB; 1WAS; 15-JAN-95.  
 DR PDB; 1WAT; 15-JAN-95.  
 DR PDB; 2LIG; 15-SEP-95.  
 DR PDB; 1LIH; 15-SEP-95.  
 DR PDB; 1VLS; 21-APR-97.  
 DR PDB; 1VLT; 15-MAY-97.  
 DR PDB; 1JMW; 22-AUG-01.  
 DR STyGene; SG10387; tar.  
 DR InterPro; IPR004091; AspME chemrecept.  
 DR InterPro; IPR004089; Chmtaxis\_transd.  
 DR InterPro; IPR003660; HAMP.  
 DR InterPro; IPR004090; Me\_chemotaxis.  
 DR InterPro; IPR003122; tarH.  
 DR Pfam; PF00672; HAMP; 1.  
 DR Pfam; PF00015; MCPsignal; 1.  
 DR PRINTS; PRO0260; CHEMTRNSDUCR.  
 DR SMART; SMO0304; HAMP; 1.  
 DR SMART; SMO0283; MA; 1.  
 DR SMART; SMO0319; TarH; 1.  
 DR PROSITE; PS00538; CHEMOTAXIS\_TRANSDUC\_1; 1.  
 DR PROSITE; PS05011; CHEMOTAXIS\_TRANSDUC\_2; 1.  
 DR PROSITE; PS0885; HAMP; 1.  
 KW Chemotaxis; Transducer; Transmembrane; Methylation; Periplasmic;  
 KW Inner membrane; 3D-structure; Complete proteome.  
 FT DOMAIN 1 6  
 FT TRANSMEM 7 33  
 FT POTENTIAL. 33  
 FT POTENTIAL. 33  
 FT DOMAIN 34 190  
 FT PERIPLASMIC (POTENTIAL). 190

FT TRANSMEM 191 211  
 FT DOMAIN 212 553  
 FT POTENTIAL. 553  
 FT TRANSMEM 212 266  
 FT DOMAIN 214 266  
 FT POTENTIAL. 266  
 FT TRANSMEM 271 500  
 FT DOMAIN 271 500  
 FT POTENTIAL. 500  
 FT TRANSMEM 64 73  
 FT SITE 64 73  
 FT POTENTIAL. 73  
 FT TRANSMEM 295 295  
 FT MOD\_RES 302 302  
 FT POTENTIAL. 302  
 FT TRANSMEM 309 309  
 FT MOD\_RES 309 309  
 FT POTENTIAL. 309  
 FT TRANSMEM 491 491  
 FT MOD\_RES 491 491  
 FT POTENTIAL. 491  
 FT TRANSMEM 451 451  
 FT CONFLICT 451 451  
 FT POTENTIAL. 451  
 FT TRANSMEM 36 38  
 FT HELIX 36 38  
 FT POTENTIAL. 38  
 FT TRANSMEM 39 73  
 FT HELIX 39 73  
 FT POTENTIAL. 73  
 FT TRANSMEM 76 77  
 FT TURN 76 77  
 FT POTENTIAL. 77  
 FT TRANSMEM 82 83  
 FT TURN 82 83  
 FT POTENTIAL. 83  
 FT TRANSMEM 89 109  
 FT TURN 89 109  
 FT POTENTIAL. 109  
 FT TRANSMEM 114 116  
 FT HELIX 114 116  
 FT POTENTIAL. 116  
 FT TRANSMEM 117 142  
 FT HELIX 117 142  
 FT POTENTIAL. 142  
 FT TRANSMEM 143 144  
 FT TURN 143 144  
 FT POTENTIAL. 144  
 FT TRANSMEM 146 151  
 FT HELIX 146 151  
 FT POTENTIAL. 151  
 FT TRANSMEM 154 177  
 FT HELIX 154 177  
 FT POTENTIAL. 177  
 FT TRANSMEM 178 179  
 FT TURN 178 179  
 FT POTENTIAL. 179  
 SQ SEQUENCE 553 AA; 59614 MW; 5CEC8E707452760D CRC64;  
 Query Match 11.3%; Score 71.5; DB 1; Length 553;  
 Best Local Similarity 28.2%; Pred. No. 16;  
 Matches 22; Conservative 15; Mismatches 28; Indels 13; Gaps 3;  
 QY 50 QQQAGARICAAQSPNDINROLVNMG-----PQQHLCNNAVEPYTSTLLFLMML--G 102  
 DB 158 QNALGEALGNVARYSENLRYQTFQSDAHYRPAQMQLG-----VLAVLVLLIMVWF 211  
 QY 103 VAGLLVGVVTVSHLSQ 120  
 DB 212 IRRALLNPLARYTIRE 229  
 RESULT 8  
 ID 34KD MYCPA STANDARD; PRT; 298 AA.  
 AC 004859;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 34 kDa antigenic protein.  
 OS Mycobacterium paratuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxId=1770;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93328703; PubMed=8335649;  
 RA Gillet P., de Kessel M., Machtelinckx L., Coene M., Cocito C.;  
 RT "Isolation and sequencing of the gene coding for an antigenic 34-  
 RT kilodalton protein of Mycobacterium paratuberculosis.";  
 RL J. Bacteriol. 175:4930-4935(1993).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV0954.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X68102; CAA48221.1; -.  
 DR PIR; A40616; A40616.  
 KW Antigen; Transmembrane.  
 FT TRANSMEM 42 62  
 FT POTENTIAL. 62

FT TRANSMEM 77 97 POTENTIAL.  
FT TRANSMEM 102 122 POTENTIAL.  
FT TRANSMEM 134 154 POTENTIAL.  
SQ SEQUENCE 298 AA; 29565 MW; B2E57AA2077F0D3D CRC64;  
Query Match 11.2%; Score 71; DB 1; Length 298;  
Best Local Similarity 27.5%; Pred. No. 9.4;  
Matches 30; Conservative 13; Mismatches 34; Indels 32; Gaps 4;  
OY 43 PGSPG-PROOQAGARLGAQSPFND-----LNROLVNMGPPQWHL-----81  
DB 5 PGSPGPPRPOSGGTATPSPFAKDDGSKPLPLYNIAVVALGPAAYLNGPFTTICA 64  
OY 82 -----GNNAVEPTISILLFLMLGVAGLL-----LYGVYIVSHS 119  
DB 65 DLPGPGIGRAGDAGTAVVALLAALLAGLLPKAKSYGVAVVAVALA 113  
RESULT 9  
GAS1\_HUMAN STANDARD; PRT; 345 AA.  
ID GAS1\_HUMAN  
AC P54826;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Growth-arrest-specific protein 1 precursor (GAS-1).  
GN GAS1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND FUNCTION.  
RC TISSUE=Liver;  
RX MEDLINE=94173926; PubMed=8127893;  
RA del Sal G., Collavin L., Ruaro M.E., Edom P., Saccone S., Valle G.D.,  
RA Schneider C.;  
RT "Structure, function, and chromosome mapping of the  
RT growth-suppressing human homologue of the murine gas1 gene";  
RL Proc Natl Acad Sci U S A. 91:1848-1852(1994).  
CC -1- FUNCTION: Specific growth arrest protein involved in growth  
CC normal and transformed cells.  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By  
CC similarity).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L13698; AAA72368.1; -;  
CC PIR; A53138;  
CC Gene; HGNC:4165; GAS1.  
CC MIM; 139185; -;  
DR GO; GO:0046658; C:extrinsic to plasma membrane, GPI-anchored; ISS.  
DR GO; GO:0007050; P:cell cycle arrest; TAS.  
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
DR GO; GO:0045749; P:negative regulation of S phase of mitotic c.; NAS.  
KW Growth arrest; GPI-anchor; Signal; Glycoprotein; Lipoprotein.  
FT SIGNAL 1 39  
FT CHAIN 40 318  
FT PROPEP 319 345  
FT DOVAIN 87 288  
FT DOVAIN 285 288  
FT DOVAIN 319 322  
FT DOVAIN 337 341  
FT LIPID 318 318  
FT CARBOHYD 117 117  
SQ SEQUENCE 345 AA; 35721 MW; 2AADS0F1D3632F9D CRC64;

Query Match 11.2%; Score 71; DB 1; Length 345;  
Best Local Similarity 29.3%; Pred. No. 11;  
Matches 27; Conservative 4; Mismatches 33; Indels 28; Gaps 2;  
OY 6 ANGEIVODDDPRVYTTTPPGPSIPROSFNKGAPGPGPPROQAGARLGAQSPFN 65  
DB 278 AGSGQPLDDDDGVHPHPPSGGAA-----SGRGDLPYFPGRRSSGGGRLAP-----326  
OY 66 DLNROLVNMGPPQWHLGNNAVEPTISILLFL 97  
DB 327 -----RGAWTPLASILLILL 341  
RESULT 10  
CTF8\_HUMAN STANDARD; PRT; 478 AA.  
ID CTF8\_HUMAN  
AC Q96D72; Q9UP86;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protein C20orf158.  
GN C20ORF158.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Bloecher H., Boecher M., Brandt P., Mewes H.-W., Gassenhuber J.,  
RA Wiemann S.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21638749; PubMed=11780052;  
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
RA Jones M., Stavaris G., Almeida J.P., Babbage A.K., Baggeley C.L.,  
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,  
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Levasainio M.H., Leverha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McCormackie L.J., McElay K., Murray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Pratchalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.,  
RA Wilmberg L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20";  
RL Nature 414:865-871 (2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Matrusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,

```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Raley U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length
RA human and mouse cDNA sequences."
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL133063; CAB61387.1; ALT_INIT.
DR EMBL; AL035669; CAC12754.1; ALT_INIT.
DR EMBL; BC012757; AAH12757.1; -.
DR Genew; HENC:16151; C200rf158.
FT DOMAIN 7 272 PRO-RICH.
FT CONFLICT 346 452 ARG-RICH.
FT SEQUENCE 478 AA; 53940 MW; EEC03CFD70BA3F00 CRC64;

Query Match 11.2%; Score 71; DB 1; Length 478;
Best Local Similarity 35.4%; Pred. No. 15;
Matches 23; Conservative 4; Mismatches 24; Indels 14; Gaps 4;

QY 10 IVDDDDPRVATTPGP-----RGSIPROSPFRNGHAGPAGP--GPRQQAARLGAAG 61
DB 24 IASNDGR-----GPPPARFAGAKGPDP--SLFSGQSPPPYGRSPSPSYLGGPAGVAP 77
QY 62 SPND 66
DB 78 SQFE 82

RESULT 11
ID CB12_MOUSE STANDARD; PRT; 481 AA.
AC Q6K3M5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE cdk5 and ab1 enzyme substrate 2 (interactor with cdk3 2) (IK3-2).
GN CAB12.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver, and Spleen;
RX MEDLINE=21952370; PubMed=11955625;
RA Sato H., Nishimoto I., Matsuo K.,
RT "Ik3-2, a relative to ik3-1/cables, is associated with cdk3, cdk5, and
RT c-abl-2";
RL Biochim. Biophys. Acta 1574:157-163(2002).
CC -!- FUNCTION: Unknown. Probably involved in G1-S cell cycle
CC transition.
CC -!- SUBUNIT: Binds to CDK3, CDK5 and c-ABL. The C-terminal cyclin-box-
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- SIMILARITY: Belongs to the cyclin family

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY049712; AAL12171.1; -.
DR MGD; MGI:2182235; Cables2.
DR GO; GO:0005515; F:protein binding; IPT.
DR InterPro; IPR006671; Cyclin_N.
DR Pfam; PF00134; cyclin; 1.
KW Cell cycle; Cell division; Cyclin.
FT DOMAIN 7 11 POLY-ALA.
FT SEQUENCE 481 AA; 52710 MW; 00DALT706578B1B2 CRC64;

Query Match 11.2%; Score 70.5; DB 1; Length 481;
Best Local Similarity 35.9%; Pred. No. 17;
Matches 23; Conservative 5; Mismatches 21; Indels 15; Gaps 4;

QY 25 PRGSPRSPFRNGHAGPAGP--GPR-RQQAARLGAAGSPNDLRQLYMNGFPQW 80
DB 269 PRPSTPR-----APPGSRHKEVPKSTPAGTELGSDGDAVEYNPLLD--DPQWP 317
QY 81 LGNH 84
DB 318 CGKH 321

RESULT 12
ID S24A_ARATH STANDARD; PRT; 1054 AA.
AC Q9SFU0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative protein transport protein Sec24-like At3g07100.
GN AT3G07100 OR T1B9.25.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoyge W., Unseld M.,
RA Pattmann B., Valle G., Bloeker H., Perez-Alonso M., Obermair B.,
RA Deisner M., Boutry M., Grivell L.A., Mache R., Paulgomech P.,
RA De Simone V., Choisme N., Artiguenave F., Robert C., Brottier P.,
RA Winkler P., Catolico L., Weissenbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Mumbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Verzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehmet T.H., Notisiek G.,
RA Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Othenaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Liauro C., Purnelle B., Masny D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Agirion A., Flores M., Liguori R., Vitale D.,
RA Mannheim G., Haase D., Schoof H., Rudi S., Zaccaria P., Wemes H.-M.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walte A., Ulteback T., Fujii C.Y., Shea T.P.,
RA Cerasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Miltischer J., Sellers P., Gill J.E., Feldlyum T.V.,
RA Preuss D., Lin X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneke T., Nakamura Y., Saito S., Kato T., Asamiy F.

```



RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;  
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis  
 thaliana.";  
 RL Mature 408,820-822(2000).  
 CC -1- FUNCTION: Component of the COP1 coat, that covers ER-derived  
 CC vesicles involved in transport from the endoplasmic reticulum to  
 CC the Golgi apparatus. COP1 is composed of at least five proteins:  
 CC the SEC23/24 complex, the SEC13/31 complex, and the protein SAR1.  
 CC Acts in the cytoplasm to promote the transport of secretory,  
 CC plasma membrane, and vacuolar proteins from the endoplasmic  
 CC reticulum to the Golgi complex (By similarity).  
 CC -1- SIMILARITY: Belongs to the SEC23/SEC24 family. SEC24 subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; AC012395; AAF20236.1; -;  
 CC DR InterPro; IPR007123; Gelsolin.  
 CC DR InterPro; IPR006900; Sec23\_helical.  
 CC DR InterPro; IPR006895; Sec23\_trunk.  
 CC DR InterPro; IPR006895; ZF-Sec23\_Sec24.  
 CC DR Pfam; PF00626; Gelsolin; 1.  
 CC DR Pfam; PF04815; Sec23\_helical; 1.  
 CC DR Pfam; PF04811; Sec23\_trunk; 1.  
 CC DR Pfam; PF04810; ZF-Sec23\_Sec24; 1.  
 CC DR Hypothetical protein; Transport; Protein transport; Golgi stack;  
 CC Endoplasmic reticulum; Multigene family.  
 CC KW DOMAIN 369 394 ZINC FINGER-LIKE.  
 CC FT DOMAIN 89 92 POLY-PRO.  
 CC FT DOMAIN 128 131 POLY-PRO.  
 CC SQ SEQUENCE 1054 AA; 115662 MW; 13FAB74147EAB9C8 CRC64;  
 QY Query Match 11.1%; Score 70; DB 1; Length 1054;  
 QY Best Local Similarity 30.7%; Pred. No. 44;  
 QY Matches 23; Conservative 6; Mismatches 32; Indels 14; Gaps 3;  
 DB 16 PRVATTPPPGSGIPROSFNFGHAPPGGPPGQQAGARLGAAGSPFNDLNRQLVNG 75  
 DB 116 PPPPTTNPFGGPPPTTL--AGHLSPMSLRPQGPAPVAVMGPPP-----GTTSG 165  
 QY 76 FPGWHLNHAVERPT 90  
 DB 166 LP---GANAYPPAT 176  
 RESULT 13  
 AANT\_HDVNA STANDARD; PRT; 214 AA.  
 AC P25880;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-MAY-1992 (Rel. 22, Last annotation update)  
 DE Delta antigen.  
 OS Hepatitis delta virus (isolate Nauru) (HDV).  
 OC Viruses; Deltaavirus.  
 OC NCBI\_TaxID=10426;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91020976; PubMed=2219700;  
 RA "Chao Y.C., Chang M.F., Gust I., Lai M.M.C.;  
 RL "Sequence conservation and divergence of hepatitis delta virus RNA.";  
 RL Virology 178:384-392(1990).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- PTM: Phosphorylated.  
 CC -1- DISEASE: BINDS SPECIFICALLY TO ANTISERA FROM PATIENTS WITH CHRONIC  
 CC HEPATITIS DELTA VIRAL INFECTIONS.

CC -1- MISCELLANEOUS: Delta antigen binds specifically to HDV RNA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; M58629; AAB59753.1; -;  
 CC DR PIR; A36212; SAVLDN.  
 CC DR InterPro; IPR002506; HDV\_ag.  
 CC DR Pfam; PF01517; HDV\_ag; 1.  
 CC DR Pfam; PD002867; HDV\_ag; 1.  
 CC KW Antigen; RNA-binding; Nuclear protein; Phosphorylation.  
 CC SQ SEQUENCE 214 AA; 24180 MW; 91F2D28B55E801D CRC64;  
 QY Query Match 10.9%; Score 69; DB 1; Length 214;  
 QY Best Local Similarity 29.3%; Pred. No. 10;  
 QY Matches 22; Conservative 8; Mismatches 33; Indels 12; Gaps 4;  
 DB 10 IVDDDFVFRITTPPPGSGIPROSFNFGHAPPGGPPGQQAGARLGAAGSPFNDLNR 69  
 DB 133 LIEEDERERRRVAGPPPGV--NPLEGSGRGAFGGQGVFVNMQ-----GVESPTRTGE 184  
 QY 70 QL---VNMGPQWHL 81  
 DB 185 GLDVRGDRGP-WDI 198  
 RESULT 14  
 IPFL\_HUMAN STANDARD; PRT; 283 AA.  
 AC P52945; O60594;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Insulin promoter factor-1 (IPF-1) (Pancreas/duodenum homeobox-1) (PDX-  
 DE 1) (Islet/duodenum homeobox-1) (IDX-1) (Somatostatin transactivating  
 DE factor-1) (STF-1) (Insulin upstream factor-1) (IUF-1) (Glucose-  
 DE sensitive factor) (GSF).  
 GN IPFL OR PDX1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96070447; PubMed=7590740;  
 RA Stoffel M., Stein R., Wright C.V., Espinosa R. III, le Beau M.M.,  
 RA Bell G.I.;  
 RT "Localization of human homeodomain transcription factor insulin  
 RT promoter factor 1 (IPF1) to chromosome band 13q12.1.";  
 RL Genomics 28:125-126(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=pancreatic islets;  
 RX MEDLINE=96220081; PubMed=8635654;  
 RA Inoue H., Riggs A.C., Tanikawa Y., Ueda K., Kuwano A., Liu L.,  
 RA Donis-Keller H., Pernutt M.A.;  
 RT "Isolation, characterization, and chromosomal mapping of the human  
 RT insulin promoter factor 1 (IPF-1) gene.";  
 RL Diabetes 45:789-794(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=pancreatic islets;  
 RA Hiroshi I.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=pancreatic islets;  
 RA Marshak S., Totary H., Cerasi E., Melloul D.;

RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Hara M., Linher T.H., Paz V.P., Wang X., Iwaseki N., Bell G.I.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Tissue-Pancreatic Islets;  
RX MEDLINE=97400587; PubMed=9252422;  
RA Macfarlane W.M., Smith S.B., James R.F., Clifton A.D., Doza Y.N.,  
RA Cohen P., Docherty K.;  
RT "The p38/activating kinase mitogen-activated protein kinase cascade  
RT mediates the activation of the transcription factor insulin upstream  
RT factor 1 and insulin gene transcription by high glucose in pancreatic  
RT beta-cells.";  
RL J. Biol. Chem. 272:20936-20944 (1997).  
RN [7]  
RP VARIANTS MODY4 ARG-18, ASN-76 AND HIS-197.  
RX MEDLINE=20015025; PubMed=10545530;  
RA Macfarlane W.M., Frayling T.M., Ellard S., Evans J.C., Allen L.I.,  
RA Sulman M.P., Ayres S., Shepherd M., Clark P., Millward A., Demaine A.,  
RA Wilkin T., Docherty K., Hattersley A.T.;  
RT "Missense mutations in the insulin promoter factor-1 gene predispose  
RT to type 2 diabetes.";  
RL J. Clin. Invest. 104:R33-R39 (1999).  
RN [8]  
RP VARIANTS MODY4 LEU-59; ASN-76 AND PRO-243 INS.  
RX MEDLINE=20015026; PubMed=10545531;  
RA Hant E.H., Scofield D.A., Chevre J.-C., Durand E., Stanojevic V.,  
RA Dina C., Habener J.F., Froguel P.;  
RT "Defective mutations in the insulin promoter factor-1 (IPF-1) gene in  
RT late-onset type 2 diabetes mellitus.";  
RL J. Clin. Invest. 104:R41-R48 (1999).  
RN [9]  
RP FUNCTION: Activates insulin, somatostatin, glucokinase, islet  
RN anyloid polypeptide and glucose transporter type 2 gene  
RN transcription. Particularly involved in glucose-dependent  
RN regulation of insulin gene transcription. Binds preferentially the  
RN DNA motif 5'-[CT]TAA[TTG]-3'. During development, specifies the  
RN early pancreatic epithelium, permitting its proliferation,  
RN branching and subsequent differentiation. At adult stage, required  
RN for maintaining the hormone-producing phenotype of the beta-cell.  
RN TCF3 (E47) and NEUROD1 and with HMG-I(Y) (By similarity).  
RN SUBUNIT: Interacts with the basic helix-loop-helix domains of  
RN TCF3 (E47) and NEUROD1 and with HMG-I(Y) (By similarity).  
RN SUBCELLULAR LOCATION: Nuclear.  
RN TISSUE SPECIFICITY: Duodenum and pancreas (Langerhans islet beta  
RN cells and small subsets of endocrine non-beta-cells; at low levels  
RN in acinar cells).  
RN DOMAIN: The Anp-type hexapeptide mediates heterodimerization of  
RN IPF1 with pax on a regulatory element of the somatostatin promoter  
RN (By similarity).  
RN DOMAIN: The homeodomain, which contains the nuclear localization  
RN signal, not only mediates DNA-binding, but also acts as a protein-  
RN protein interaction domain for TCF3 (E47), NEUROD1 and HMG-I(Y) (By  
RN similarity).  
RN PPM: PHOSPHORYLATED BY THE SAPK2 PATHWAY AT HIGH GLUCOSE  
RN CONCENTRATION.  
RN DISEASE: Defects in IPF1 are a cause of pancreatic agenesis  
RN (MIM:260370). This autosomal recessive disorder is characterized  
RN by absence or hypoplasia of pancreas, leading to early-onset  
RN insulin-dependent diabetes mellitus. This was found in a  
RN frameshift mutation that produces a truncated protein and results  
RN in a second initiation that produces a second protein that act as  
RN a dominant negative mutant.  
RN DISEASE: Defects in IPF1 are a cause of susceptibility to  
RN noninsulin-dependent diabetes mellitus (NIDDM) (MIM:125853); also  
RN known as diabetes mellitus type II.  
RN DISEASE: Defects in IPF1 are the cause of maturity onset diabetes  
RN of the young type IV (MODY4) (MIM:606392); also symbolized MODY-  
RN IV. MODY4 is a form of non-insulin-dependent diabetes mellitus  
RN (NIDDM) characterized by an autosomal dominant mode of  
RN inheritance, age of onset of 25 years or younger and a primary  
RN defect in insulin secretion.

CC IPF1/XLHBOX8 SUBFAMILY.  
CC -1- SIMILARITY: Contains 1 homeobox domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----  
CC EMBL: U35632; AAA88820.1; -;  
CC EMBL: S82178; AAB47101.1; -;  
CC EMBL: S82168; AAB47101.1; JOINED.  
CC EMBL: U30329; AAB47012.1; -;  
CC EMBL: X99894; CAA68169.1; -;  
CC EMBL: AF035260; AAB88463.1; -;  
CC EMBL: AF035258; AAB88463.1; JOINED.  
CC EMBL: AF049893; AAC05157.1; -;  
CC PIR: G01926; G01926.  
CC HSSP: P02833; 9ANT.  
CC TRANSFAC: T04362; -;  
CC Gene: HGNC:6107; IPF1.  
CC MIM: 600733; -;  
CC MIM: 260370; -;  
CC MIM: 125853; -;  
CC MIM: 606392; -;  
CC GO: GO:0003704; F:specific RNA polymerase II transcription fa. .; TAS.  
CC GO: GO:0006091; P:energy pathways; TAS.  
CC GO: GO:0007397; P:histogenesis and organogenesis; TAS.  
CC InterPro: IPR001827; Antennapedia.  
CC InterPro: IPR001356; Homeobox.  
CC InterPro: IPR000047; HTH\_LambdaRepressor.  
CC Pfam: PF00046; homeobox; 1.  
CC PRINTS: PR00025; ANTENNAPEDIA.  
CC PRINTS: PR00024; HOMEBOX.  
CC PRINTS: PR00031; HTRREPRESSR.  
CC ProDom: PD000010; Homeobox; 1.  
CC SMART: SM00388; HOX; 1.  
CC PROSITE: PS00027; HOMEBOX\_1; 1.  
CC PROSITE: PS00021; HOMEBOX\_2; 1.  
CC PROSITE: PS00032; ANTENNAPEDIA; FALSE NEG.  
CC KX Transcription regulation; Activator; DNA-binding; Homeobox;  
CC KW Nuclear protein; Developmental protein; Disease mutation;  
CC KW Diabetes mellitus; Phosphorylation; Polymorphism.  
CC FT DOMAIN 13 73 TRANSACTIVATION DOMAIN (BY SIMILARITY).  
CC FT SITE 42 51 POLY-PRO.  
CC FT SITE 118 123 ANTP-TYPE HEXAPEPTIDE.  
CC FT DNA\_BIND 146 205 HOMEBOX.  
CC FT DOMAIN 197 203 NUCLEAR LOCALIZATION SIGNAL (BY  
CC FT DOMAIN 216 219 SIMILARITY).  
CC FT DOMAIN 239 244 POLY-GLY.  
CC FT VARIANT 18 18 C->R (IN MODY4).  
CC FT VARIANT 59 59 Q->L (IN MODY4).  
CC FT VARIANT 76 76 D->N (IN MODY4; COULD BE A  
CC FT VARIANT 76 76 POLYMORPHISM).  
CC FT VARIANT 197 197 /FTID=VAR\_009311.  
CC FT VARIANT 243 243 R->H (IN MODY4).  
CC FT VARIANT 243 243 /FTID=VAR\_009312.  
CC FT VARIANT 243 243 P->P (IN MODY4).  
CC FT VARIANT 243 243 /FTID=VAR\_009313.  
CC FT CONFLICT 56 56 A->S (IN REF. 6).  
CC FT CONFLICT 116 116 Q->H (IN REF. 6).  
CC FT CONFLICT 210 211 GG->SS (IN REF. 6).  
CC SQ SEQUENCE 283 AA; 30771 MW; 5048EC8D3289F72B CRC64;  
Query Match 10.9%; Score 69; DB 1; Length 283;  
Best Local Similarity 27.7%; Pred. No. 14;  
Matches 28; Conservative 7; Mismatches 38; Indels 28; Gaps 5;

QY 6 ANGEIVQDDPRVTRTTQPPRGSIIPROSPFNRHGA-----PPGGPGRQQAAGATGA 59  
 DB 53 ALGLEQSPDPDISPYEVPFLAEDPAVAHLHHLPAQLALPHRPAAGFPF-----GAAPGV 108  
 QY 60 AQSPPNDLNRLQVYMGFP-----QWHLGNHAYEP 88  
 DB 109 LEEP-----NR--VOLPFPWMKSTKAHAWKQWAGAYAAEP 143

RESULT 15  
 ID IPFL\_MESAU STANDARD; PRT; 283 AA.  
 AC P70118;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Insulin promoter factor 1 (IPF-1) (Homeobox domain protein PDX1).  
 GN IPFL OR PDX1.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rudnick A., Ling T.Y., Odagiri H., Rutter W.J., German M.S.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Activates insulin and somatostatin gene transcription.  
 CC Key regulator of islet peptide hormone expression but also  
 CC responsible for the development of the pancreas, most probably by  
 CC determining maturation and differentiation of common pancreatic  
 CC precursor cells in the developing gut. Binds the DNA sequence 5'-  
 CC CC[CT]TATGGG-3' (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -1- TISSUE SPECIFICITY: Pancreas; Islet beta-cells (By similarity).  
 CC -1- DOMAIN: The Anp-type hexapeptide mediates heterodimerization of  
 CC IPFL with PRX on a regulatory element of the somatostatin promoter  
 CC (By similarity).  
 CC -1- DOMAIN: The homeobox domain, which contains the nuclear localization  
 CC signal, not only mediates DNA-binding, but also acts as a protein-  
 CC protein interaction domain for TCF3 (E47), NEUROD1 and HMG-I(Y) (By  
 CC similarity).  
 CC -1- SIMILARITY: BELONGS TO THE ANTP HOMEBOX FAMILY.  
 CC IPFL/XLHBOX SUPERFAMILY.  
 CC -1- SIMILARITY: Contains 1 homeobox domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U73854; AAB18252.1; -.  
 DR HSSP; P02833; 9ANT.  
 DR TRANSPAC; T04267; -.  
 DR InterPro; IPR001827; Antennapedia.  
 DR InterPro; IPR001355; Homeobox.  
 DR InterPro; IPR000047; HTH\_1amdrepressr.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00025; ANTENNAPEDIA.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOK; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS00071; HOMEBOX\_2; 1.  
 DR PROSITE; PS00032; ANTENNAPEDIA; FALSE NEG.  
 KM Transcription regulation; Activator; DNA-binding; Homeobox;  
 KM Nuclear protein.  
 FT DOMAIN 13 73 TRANSACTIVATION DOMAIN (BY SIMILARITY).  
 FT 42 48 POLY-PRO.

FT	SITE	118	123	ANTP-TYPE HEXAPEPTIDE.
FT	DNA BIND	146	205	HOMEBOX.
FT	DOMAIN	197	203	NUCLEAR LOCALIZATION SIGNAL (BY
FT	DOMAIN	239	244	SIMILARITY).
FT	SEQUENCE	283 AA;	30819 MW;	POLY-PRO.
SO	SEQUENCE	283 AA;	30819 MW;	22714BBSARD54BF2 CRC64;

Query Match 10.9%; Score 69; DB 1; Length 283;  
 Best Local Similarity 26.4%; Pred. No. 14;  
 Matches 24; Conservative 8; Mismatches 51; Indels 8; Gaps 2;

QY 6 ANGEIVQDDPRVTRTTQPPRGSIIPROSPFNR-----GHGAPPGGPGGPRQQAAG--RL 57  
 DB 53 ALGLEQSPDPDISPYEVPFLAEDPAVAHLHHLPAQLALPHRPAAGFPFNGTEPGGLEEP 112  
 QY 58 AQSPPNDLNRLQVYMGFPQWHLGNHAYEP 88  
 DB 113 SRQQLPFPWMKSTKAHAWKQWAGAYAAEP 143

Search completed: August 24, 2004, 09:18:24  
 Job time : 12 secs

Blank



ID Q9D882 PRELIMINARY; PRT; 120 AA.  
 AC Q9D882;  
 DT 01-JUN-2001 (TRENBLREL. 17, Created)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update) (Hypothetical  
 DE 2010107G23Rik protein (RIKEN CDNA 2010107G23 gene)  
 DE protein).  
 GN 2010107G23Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Small intestine;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;  
 RX MEDLINE=22354683; PubMed=12468551;  
 RA The FANTOM Consortium,  
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK008338; BAB25613.1; -  
 DR EMBL; BC024943; AAH24943.1; -  
 DR EMBL; AK035688; BAC29152.1; -  
 DR MGD; MGI:1917144; 2010107G23Rik.  
 KW Hypothetical protein.  
 SQ SEQUENCE 120 AA; 13118 MW; 92C337D6CC6B4460 CRC64;

Query Match 90.0%; Score 568.5; DB 11; Length 120;  
 Best Local Similarity 92.6%; Pred. No. 7.8e-54;  
 Matches 112; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

ID Q9CZL2 PRELIMINARY; PRT; 131 AA.  
 AC Q9CZL2;  
 DT 01-JUN-2001 (TRENBLREL. 17, Created)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)  
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)  
 DE 270063A17Rik protein.  
 DE 5730508B09Rik OR 270063A17Rik.  
 GN 5730508B09Rik OR 270063A17Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCB1\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 RL EMBL; AK012470; BAB28264.1; -  
 DR MGD; MGI:1917867; 5730508B09Rik.  
 SQ SEQUENCE 131 AA; 14511 MW; 2F4937AADF6F169 CRC64;

Query Match 18.8%; Score 118.5; DB 11; Length 131;  
 Best Local Similarity 28.6%; Pred. No. 5.3e-05;  
 Matches 34; Conservative 19; Mismatches 49; Indels 17; Gaps 3;

ID Q9N8U7 PRELIMINARY; PRT; 132 AA.  
 AC Q9N8U7;  
 DT 01-OCT-2002 (TRENBLREL. 22, Created)  
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)  
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)  
 DE Hypothetical protein FLJ39370.  
 DE Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,  
 RA Arita M., Mutsaers K., Yuuki H., Hara H., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto U., Isono Y.,  
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,

RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,  
 RA Sugano S., Nagahari K., Masubo Y., Nagai K., Isogai T.,  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK096689; BAC04841.1; -  
 KM Hypothetical protein.  
 SQ SEQUENCE 132 AA; 14725 MW; D241F43200E41FDE CRC64;  
 Query Match 17.5%; Score 110.5; DB 4; Length 132;  
 Best Local Similarity 31.9%; Pred. No. 0.00039;  
 Matches 30; Conservative 17; Mismatches 34; Indels 13; Gaps 3;  
 QY 26 RGSIPRQS---FNRHGHAPPGSGPFGPQQAGARLGAAGSPFNDLNRQLVNMGPQWHL 81  
 DB 42 RGSIPRQS---FNRHGHAPPGSGPFGPQQAGARLGAAGSPFNDLNRQLVNMGPQWHL 81  
 QY 82 GNHAEVPTSLILFLMMGLVGRGLLVGLVYL 115  
 DB 93 GERIVEPVIVIFFWMLMFLGLQALGLVAVLCV 126  
 RESULT 5  
 O8TRF9 PRELIMINARY; PRT; 132 AA.  
 AC O8TRF9;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Hypothetical protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC022534; AAR22534.1; -  
 KM Hypothetical protein.  
 SQ SEQUENCE 132 AA; 14787 MW; 284B5499A1E50F05 CRC64;  
 Query Match 17.4%; Score 110; DB 4; Length 132;  
 Best Local Similarity 29.3%; Pred. No. 0.00044;  
 Matches 27; Conservative 14; Mismatches 29; Indels 22; Gaps 1;  
 QY 46 PGRPOQAGARLGAAGSP-----FNDLNRQLVNMGPQWHL 83  
 DB 35 PGRPOQAGARLGAAGSP-----FNDLNRQLVNMGPQWHL 83  
 QY 84 HAVPVTSLILFLMMGLVGRGLLVGLVYL 115  
 DB 95 RIVERPVIVIFFWMLMFLGLQALGLVAVLCV 126  
 RESULT 6  
 O82HR5 PRELIMINARY; PRT; 223 AA.  
 AC O82HR5;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 NC NCB1\_TaxID=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis."  
 RL Nat. Biotechnol. 21:526-531 (2003).  
 DR EMBL; AP005035; BAC71155.1; -  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 223 AA; 23128 MW; AF13F56F4DB3A87 CRC64;  
 Query Match 15.7%; Score 99.5; DB 16; Length 223;  
 Best Local Similarity 30.9%; Pred. No. 0.011;  
 Matches 29; Conservative 9; Mismatches 21; Indels 35; Gaps 4;  
 QY 23 QPP---RGSIPRQS---FNRHGHAP--PGSGPFGPQQAGARLGAAGSPFNDLNRQL 71  
 DB 27 QPPYQAPRGIPIQGGIGYIPQGGPQGYIPGGPFGPQVASKGRRLGA----- 74  
 QY 72 VNMGPQWHLGNHAEVPTSLILFLMMGLVGRGL 105  
 DB 75 -----RAIDGVAFFVIFLGLAGVAG 96  
 RESULT 7  
 O9NT75 PRELIMINARY; PRT; 219 AA.  
 AC O9NT75;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 GN DKFZP434N2030.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL137488; CAB70763.1; -  
 DR PIR; T46472; T46472.  
 KM Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 219 AA; 24989 MW; EB4D7AC00C28F97B CRC64;  
 Query Match 13.2%; Score 83.5; DB 4; Length 219;  
 Best Local Similarity 36.5%; Pred. No. 0.6;  
 Matches 23; Conservative 11; Mismatches 24; Indels 5; Gaps 2;  
 QY 41 APPGPGPQQAGARLGAAGSPFNDLNRQLVNMGPQWHLGNHAEVPTSLILFLMM 100  
 DB 9 APPRAGSVRRTRRG-RGLHLEIPC-----RPGISVGFPGSVWGGSVRELVVLFVYR 63  
 QY 101 LGV 103  
 DB 64 EGI 66  
 RESULT 8  
 O7XRN3 PRELIMINARY; PRT; 174 AA.  
 AC O7XRN3;  
 ID O7XRN3;  
 AC O7XRN3;

DT 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE OSJNBA0024J22.19, protein.  
GN OSJNBA0024J22.19.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Echaritoidae; Oryzae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Wang Q.J., Zhang L.,  
RA Lu Y.Q., Yu S.L., Liu X.H., Li T.T., Zhang Y.S., Lu Y., Li C., Li T.,  
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,  
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,  
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,  
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.,  
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AL731596; CA02411.1; -  
SQ SEQUENCE 174 AA; 17749 MW; 6BD95E3BD5117DC7 CRC64;

Query Match 13.1%; Score 83; DB 10; Length 174;  
Best Local Similarity 32.1%; Pred. No. 0.52; Mismatches 23; Indels 24; Gaps 3;  
Matches 27; Conservative 10;

QY 38 GHGAPGPGPQQAGARLGAQSPFNDLRQLVNMGFQWLGNAVEPVT----- 90  
DB 14 GCGAPPTGAG-----SGAGAGAPGAGVNDNN-----CGNSAPSSGGLPSGY 56  
QY 91 SILLFLMLTGVKGLIVGLYL 114  
DB 57 NLLFLFLALFLPPLALSLYYV 80

RESULT 9  
ID 036094 PRELIMINARY; PRT; 102 AA.  
AC 036094;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hepatitis delta antigen (Fragment).  
OS Hepatitis delta virus (HDV).  
OS Viruses; Deltavirus.  
OX NCBI\_TaxID=12475;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=Turkish isolate 07;  
RX MEDLINE=97378276; PubMed=9234957;  
RA Shail A.O., Haddiyannis S., Hootnagle J.H., Di Bisceglie A.M.,  
RA Gerin J.L., Casey J.L.;  
RT "Geographic distribution and genetic variability of hepatitis delta  
RT virus genotype 1.";  
RL Virology 234:160-167(1997).  
DR EMBL; AF008349; AAB66964.1; -  
DR GO; GO:0042025; C:host cell nucleus; IEA.  
DR GO; GO:0003723; F:rRNA binding; IEA.  
DR InterPro; IPR002506; HDV\_ag;  
DR Pfam; PF01517; HDV\_ag; 1.  
DR Prodom; PD002887; HDV\_ag; 1.  
FT NON TER 1  
SQ SEQUENCE 102 AA; 11085 MW; 3136BB4EA39F792 CRC64;

Query Match 13.0%; Score 82; DB 12; Length 102;  
Best Local Similarity 32.9%; Pred. No. 0.36; Mismatches 27; Indels 14; Gaps 5;  
Matches 25; Conservative 10;

DB 21 LITEDERRERRTAGPPTGVNPSB---GERRAGGGGFVPSKQ-----GVPSFPHRLG 71  
QY 69 RQL---VVMGFQWHL 81  
DB 72 EGLDVRSQGF-WDI 86

RESULT 10  
ID 0905K9 PRELIMINARY; PRT; 608 AA.  
AC 0905K9;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE NTR.  
OS Herpesvirus papio.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gammapherpesvirinae; Lymphocryptovirine.  
OX NCBI\_TaxID=10394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=Baboon lymphocryptovirus BA65;  
RA Zeng J.-C., Ryan J., Ling F.D., Loeb D.D., Pagano J.S., Hayward S.D.,  
RA Hayward G.S.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF200364; AAF23950.1; -  
SQ SEQUENCE 608 AA; 60458 MW; 4BF82ACB0A029627 CRC64;

Query Match 13.0%; Score 82; DB 12; Length 608;  
Best Local Similarity 47.4%; Pred. No. 2.8; Mismatches 17; Indels 0; Gaps 0;  
Matches 18; Conservative 3;

QY 24 PPRGSIIPRSGPFRNGHAPGPGPQQAGARLGAQ 61  
DB 65 PPGAGQRPSPGPTGCHPAAPGAPRSPRTRRRGSAG 102

RESULT 11  
ID 09A913 PRELIMINARY; PRT; 475 AA.  
AC 09A913;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Sensor histidine kinase.  
GN CCL181.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;  
OC Caulobacteraceae; Caulobacter.  
OX NCBI\_TaxID=155892;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=ATCC 19089 / CBL5;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nieman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Eisen J., Heidelberg J.F., Alley M.R., Chua N., Maddock J.R.,  
RA Potocka I., Nelson W.C., Newton R.J., Durkin A.S., Gwinn M.L., Hatt D.H.,  
RA DeBoy R.T., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Kolonay J.F., Smit J., Wolf A., Vamathavan J., Ernoiaeva M., White O.,  
RA Ueberback T., Tran K., Wolf A., Vamathavan J., Ernoiaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of Caulobacter crescentus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
CC -1- SIMILARITY: TO PROKARYOTE SENSOR TRANSDUCTION PROTEINS.  
DR EMBL; AE005795; AAK23165.1; -  
DR PIR; A87396; A87396.  
DR HSSP; P02933; 1BXD.  
DR TIGR; CCL181; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0016501; F:kinase activity; IEA. TT







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:17:06 ; Search time 19 Seconds

(without alignments)  
328.776 Million cell updates/sec

Title: US-10-001-885-125

Perfect score: 632 1 MRRILANGSLVQDDPRVPT.....GVAGLLLVGLVIVLSQR 121

Sequence:

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

389414

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/1aa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/1aa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/1aa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/1aa/PTCUTS COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/1aa/backfilltest.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	551	87.2	104	US-09-621-976-4724 Sequence 4724, App
2	387	61.2	111	US-08-149-476-673 Sequence 673, App
3	245	38.8	113	US-09-149-476-673 Sequence 367, App
4	240	38.0	113	US-09-621-976-6924 Sequence 6924, App
5	88	13.9	460	US-09-489-039A-8966 Sequence 8966, App
6	79.5	12.6	763	US-09-252-991A-30146 Sequence 30146, App
7	78.5	12.4	605	US-08-889-402-2 Sequence 2, App1
8	76	12.0	217	US-09-252-991A-23015 Sequence 23015, App
9	75	11.9	200	US-09-252-991A-22303 Sequence 22303, App
10	75	11.9	655	US-09-252-991A-25314 Sequence 25314, App
11	74.5	11.8	215	US-09-252-991A-31559 Sequence 31559, App
12	74	11.7	263	US-09-252-991A-28276 Sequence 28276, App
13	73.5	11.6	638	US-09-252-991A-24325 Sequence 24325, App
14	73	11.6	270	US-09-252-991A-22026 Sequence 22026, App
15	73	11.6	351	US-09-252-991A-32396 Sequence 32396, App
16	72.5	11.5	482	US-09-252-991A-31705 Sequence 31705, App
17	71.5	11.3	246	US-09-252-991A-28637 Sequence 28637, App
18	71.5	11.3	415	US-09-252-991A-23751 Sequence 23751, App
19	71.5	11.3	427	US-09-252-991A-31904 Sequence 31904, App
20	71.5	11.3	734	US-09-252-991A-33036 Sequence 33036, App
21	71.5	11.3	1053	US-09-252-991A-26140 Sequence 26140, App
22	71	11.2	199	US-08-122-458D-18 Sequence 18, App1
23	71	11.2	298	US-08-122-458D-11 Sequence 11, App1
24	71	11.2	480	US-09-252-991A-23424 Sequence 23424, App
25	71	11.2	645	US-09-252-991A-22095 Sequence 22095, App
26	71	11.2	988	US-09-252-991A-29659 Sequence 29659, App
27	70.5	11.2	149	US-09-252-991A-22850 Sequence 22850, App

28	70.5	11.2	295	US-09-252-991A-22439	Sequence 22439, A
29	70.5	11.2	390	US-09-252-991A-22732	Sequence 22732, A
30	70.5	11.2	531	US-09-252-991A-24404	Sequence 24404, A
31	70.5	11.2	863	US-09-252-991A-21831	Sequence 21831, A
32	70	11.1	253	US-09-252-991A-27487	Sequence 27487, A
33	70	11.1	380	US-09-252-991A-22319	Sequence 22319, A
34	69.5	11.0	324	US-09-252-991A-24165	Sequence 24165, A
35	69.5	11.0	329	US-09-252-991A-18860	Sequence 18860, A
36	69.5	11.0	410	US-09-252-991A-25451	Sequence 25451, A
37	69.5	11.0	526	US-09-252-991A-21748	Sequence 21748, A
38	69.5	11.0	579	US-09-252-991A-30344	Sequence 30344, A
39	69	10.9	243	US-09-252-991A-25431	Sequence 25431, A
40	69	10.9	478	US-09-252-991A-24169	Sequence 24169, A
41	69	10.9	871	US-09-252-991A-19431	Sequence 19431, A
42	68.5	10.8	225	US-09-252-991A-16928	Sequence 16928, A
43	68.5	10.8	305	US-09-252-991A-27334	Sequence 27334, A
44	68.5	10.8	406	US-09-489-039A-12736	Sequence 12736, A
45	68.5	10.8	555	US-09-252-991A-31521	Sequence 31521, A

## ALIGNMENTS

RESULT 1  
US-09-621-976-4724 Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: J. Y.  
FILE OF INVENTION: ESTs and Encoded Human Proteins:  
TITLE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO: 4724  
LENGTH: 104  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-621-976-4724

Query Match 87.2%; Score 551; DB 4; Length 104;  
Best Local Similarity 100.0%; Pred. No. 7.4e-57;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRRILANGSLVQDDPRVPTTTPRGSTPRGFFNRGAPGPGPQQAGARLGA	60
DB	1	MRRILANGSLVQDDPRVPTTTPRGSTPRGFFNRGAPGPGPQQAGARLGA	60
QY	61	QSPFNDLNQLVNMGFPQWHLGNHVEPVTSILLFLMLGVR	104
DB	61	QSPFNDLNQLVNMGFPQWHLGNHVEPVTSILLFLMLGVR	104

RESULT 2  
US-09-149-476-673 Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
FILE OF INVENTION: 186 Human Secreted proteins  
TITLE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-09-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621

[illegible]

EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 61.2%; Score 387; DB 4; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1,1e-37;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLTNGEIVQDDPRVTRTTTPPGSIPROSFNRRGAGPAGGPPROQQAARLGA 60  
DB 39 MVRILNGEIVQDDPRVTRTTTPPGSIPROSFNRRGAGPAGGPPROQQAARLGA 98  
QY 61 QSPFNDRQLVN 73  
DB 99 QSPFNDRQLVN 111

RESULT 3  
US-09-149-476-367  
Sequence 367, Application US/09149476  
Patent No. 6420526  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/09/149,476  
CURRENT FILING DATE: 1998-03-08  
EARLIER APPLICATION NUMBER: PCT/US98/04493  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,500  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,613  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,596  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,601  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315

EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

## Query Match

Best Local Similarity 38.8%; Score 245; DB 4; Length 48;  
Pred. No. 1,3e-21; Indels 0; Gaps 0;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 MGFPMHGNHNAVEPVTSILLLFILMPLGVRLVGLVIVSHLSOR 121  
Db 1 MGFPMHGNHNAVEPVTSILLLFILMPLGVRLVGLVIVSHLSOR 48

## RESULT 4

US-09-621-976-6924  
Sequence 6924, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
NUMBER OF SEQ ID NOS: 2000-07-21  
SOFTWARE: Patent.pm  
SEQ ID NO 6924  
LENGTH: 113  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: 97  
OTHER INFORMATION: Xaa = Pro,Gln  
US-09-621-976-6924

Query Match 38.8%; Score 240; DB 4; Length 113;  
Best Local Similarity 83.9%; Pred. No. 1.6e-20;  
Matches 47; Conservative 1; Mismatches 6; Indels 2; Gaps 1;

QY 1 MVRILANGELVQDDPRVTTQPPRGSIIPROSGFNRGHGAPG--GPGRQOQAG 54  
DB 43 MVRILANGELVQDDPRVTTQPPRGSIIPROSGFNRGHGAPGVLA-PASSRQVXG 98

RESULT 5  
US-09-489-039A-8966  
; Sequence 8966, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709,2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 8966  
; LENGTH: 460  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-8966

Query Match 13.9%; Score 88; DB 4; Length 460;  
Best Local Similarity 31.4%; Pred. No. 0.052;  
Matches 22; Conservative 17; Mismatches 25; Indels 6; Gaps 3;  
QY 56 RLGAAQSPNDLNQVYM-GFQPMHGNRAVEPTVTLILFLMLGVAGLLNGVYL 114  
DB 213 RLAVASBPQDGINMIAVLIGIPMHVGTLPKMTISVILIGLILGDAIV-LVFI 271  
QY 115 ----VSHLSQ 120  
DB 272 QPVLNSMSE 281

RESULT 6  
US-09-252-991A-30146  
; Sequence 30146, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196,136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30146  
; LENGTH: 763  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30146

Query Match 12.6%; Score 79.5; DB 4; Length 763;  
Best Local Similarity 39.0%; Pred. No. 0.99;  
Matches 23; Conservative 5; Mismatches 22; Indels 9; Gaps 3;  
QY 12 QDDDPVRYTTQPPRGSIIPROSGFNRGHG-----PFG--PGRQOQAGRLGAASPFNDLNRLV 63  
DB 443 QGGDGRPRTPDERAGATPGQAL--RAQGA-PAGTLPPGGRAGPRRROPAGGRBAGPEHP 499

RESULT 7  
US-08-889-402-2  
; Sequence 2, Application US/08889402

Patent No. 5811288  
; GENERAL INFORMATION:  
; TITLE OF INVENTION: NOVEL ISOFORM GENE FOR FOCAL ADHESION  
; TITLE OF INVENTION: PROTEIN PAXILIN  
; NUMBER OF SEQUENCES: 7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/889,402  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 178334/1996  
; FILING DATE: 08-JUL-1996  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 605 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; CELL TYPE: Monoblastic cell and placenta  
; CELL LINE: U937  
US-08-889-402-2

Query Match 12.4%; Score 78.5; DB 2; Length 605;  
Best Local Similarity 34.9%; Pred. No. 0.96;  
Matches 22; Conservative 5; Mismatches 29; Indels 7; Gaps 2;  
QY 16 PVRITTPPPRGSIIPROSGF--NRGHGAPGPGSPRQOQAGARLGAASPFNDLNRLV 72  
DB 312 PSLRSPDQPF-----PQPMAGAKTGSSPFGPPKPSQDLSMGLSDJNKGVATV 367  
QY 73 NMG 75  
DB 368 AKG 370

RESULT 8  
US-09-252-991A-23015  
; Sequence 23015, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196,136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23015  
; LENGTH: 217  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23015

Query Match 12.0%; Score 76; DB 4; Length 217;  
Best Local Similarity 32.9%; Pred. No. 0.49;  
Matches 23; Conservative 3; Mismatches 18; Indels 26; Gaps 4;  
QY 26 RGSIPROSGFNRGHGAPG-----GP-----GPRQOQAGARLGAASPFNDLNRLV 72  
DB 91 RAGVPRARS-RRRGHRPRQMVADRAGFTAGLPAQPRLHDAGASLPGSDPF----- 140  
QY 73 NMGPFQMHG 82

Db 141 ---GPAWRG 147

## RESULT 9

US-09-252-991A-22303  
; Sequence 22303, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 22303  
; LENGTH: 200  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-22303

Query Match 11.9%; Score 75; DB 4; Length 200;  
Best Local Similarity 33.3%; Pred. No. 0.58; 29; Indels 10; Gaps 2;  
Matches 22; Conservative 5; Mismatches 29;

Qy 15 DPRVTTTPPGSGIPRGSFFNRHGAPGPGPGROOAGARLGAQSPFNDLRLVNM 74  
Db 40 DEHYRRSQHPRRAKRCATV---HRRPGEPPDLPQVGFRAAARP-----RLAGL 89  
Qy 75 GFGQNH 80  
Db 90 GVQGRH 95

## RESULT 10

US-09-252-991A-25314  
; Sequence 25314, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25314  
; LENGTH: 655  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25314

Query Match 11.9%; Score 75; DB 4; Length 655;  
Best Local Similarity 33.3%; Pred. No. 2.7; 30; Indels 20; Gaps 6;  
Matches 29; Conservative 8; Mismatches 30;

Qy 3 RILANGELIVQ-DDDPR---VTTTTPPGSGIPRGSFFNRHGAPGPGPGROOAGARL 57  
Db 516 RIVEGGEVMEVGDPRHARHPKVPVQPAR-----RPPGRPPGAPAGSHR--RRP 563  
Qy 58 GAQSPFNDLRLVNMGFPQMHGNNH 84  
Db 564 GAARSPVH--RRQPTGAG-PAGHNGSH 587

## RESULT 11

US-09-252-991A-31559  
; Sequence 31559, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31559  
; LENGTH: 215  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31559

Query Match 11.8%; Score 74.5; DB 4; Length 215;  
Best Local Similarity 40.9%; Pred. No. 0.73; 20; Indels 1; Gaps 1;  
Matches 18; Conservative 5; Mismatches 20;

Qy 23 OPPRGSIPRGSFFNRHGAPG-GPGPGROOAGARLGAQSPFN 65  
Db 98 RPPRALQSRAGRAARRRPPVLPGRRAQLGCRPGAGQNNFN 141

## RESULT 12

US-09-252-991A-28276  
; Sequence 28276, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28276  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28276

Query Match 11.7%; Score 74; DB 4; Length 263;  
Best Local Similarity 41.7%; Pred. No. 1.1; 7; Indels 22; Gaps 6;  
Matches 25; Conservative 6; Mismatches 7;

## RESULT 13

US-09-252-991A-24325  
; Sequence 24325, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24325  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24325

Qy 15 DPRVTTTTPPGSGIPRGSFFNRHGAPGPGPGROOAGARLGAQSP 63  
Db 110 DPRLR---QPVDAQDPAQDPR-----RGHGA---GPGFWLEPPRRRTAQAQAPAEHP 158





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2004, 09:19:19 ; Search time 124 Seconds  
(without alignments)  
306.652 Million cell updates/sec

Title: US-10-001-885-125

Sequence: 1 MRLIANGELIVQDDPRVRR.....GVGGLLVGLVIVSHLSQR 121

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1295152 seqs, 31425058 residues

Total number of hits satisfying chosen parameters: 1295152

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgml\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*
- 2: /cgml\_6/ptodata/1/pubppaa/PCRN\_NEW\_PUB.pep.\*
- 3: /cgml\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*
- 4: /cgml\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*
- 5: /cgml\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*
- 6: /cgml\_6/ptodata/1/pubppaa/PCRN\_NEW\_PUB.pep.\*
- 7: /cgml\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*
- 8: /cgml\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
- 9: /cgml\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
- 10: /cgml\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
- 11: /cgml\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
- 12: /cgml\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep.\*
- 13: /cgml\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
- 14: /cgml\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
- 15: /cgml\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
- 16: /cgml\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep.\*
- 17: /cgml\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*
- 18: /cgml\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	632	100.0	121	US-10-001-885-125	Sequence 125, App
2	632	100.0	121	US-10-076-747-132	Sequence 122, App
3	387	61.2	111	US-09-809-391-673	Sequence 673, App
4	387	61.2	111	US-09-882-171-673	Sequence 673, App
5	387	61.2	111	US-10-164-861-673	Sequence 673, App
6	245	38.8	48	US-09-809-391-367	Sequence 367, App
7	245	38.8	48	US-09-882-171-367	Sequence 367, App
8	245	38.8	48	US-10-164-861-367	Sequence 367, App
9	110.5	17.5	132	US-10-108-260A-3882	Sequence 3882, App
10	99.5	15.7	223	US-10-156-761-1980	Sequence 1980, App
11	85.5	13.5	112	US-10-424-599-174474	Sequence 174474, App
12	82	13.0	123	US-10-437-963-198292	Sequence 198292, App
13	81	12.8	256	US-10-425-114-60982	Sequence 60982, App
14	80	12.7	92	US-10-106-698-8371	Sequence 8371, App
15	80	12.7	392	US-10-156-761-10160	Sequence 10160, App

16	80	12.7	542	16	US-10-437-963-145131	Sequence 145131, App
17	79.5	12.6	286	16	US-10-437-963-187696	Sequence 187696, App
18	76.5	12.1	225	16	US-10-437-963-177234	Sequence 177234, App
19	76.5	12.1	224	16	US-10-437-963-164016	Sequence 164016, App
20	76	12.0	331	16	US-10-425-114-64460	Sequence 64460, App
21	75.5	11.9	622	15	US-10-240-145-133	Sequence 133, App
22	75	11.9	701	15	US-10-108-260A-2701	Sequence 2701, App
23	73.5	11.6	236	16	US-10-437-963-149016	Sequence 149016, App
24	72.5	11.5	137	16	US-10-437-963-193889	Sequence 193889, App
25	72.5	11.5	440	15	US-10-369-493-18063	Sequence 18063, App
26	72	11.4	259	12	US-10-425-114-70983	Sequence 70983, App
27	72	11.4	1126	16	US-10-437-963-183078	Sequence 183078, App
28	71.5	11.3	195	14	US-10-029-386-32263	Sequence 32263, App
29	71.5	11.3	279	12	US-10-425-114-40428	Sequence 40428, App
30	71.5	11.3	378	16	US-10-437-963-150350	Sequence 150350, App
31	71.5	11.3	1336	16	US-10-437-963-115864	Sequence 115864, App
32	71.5	11.3	19662	15	US-10-084-846A-6	Sequence 6, App
33	71	11.2	166	14	US-10-156-761-10935	Sequence 10935, App
34	71	11.2	642	12	US-10-263-929-105	Sequence 105, App
35	71	11.2	793	14	US-10-029-386-32918	Sequence 32918, App
36	70.5	11.2	140	12	US-10-424-599-231854	Sequence 231854, App
37	70.5	11.2	248	12	US-10-425-114-42547	Sequence 42547, App
38	70.5	11.2	915	13	US-10-029-386-156444	Sequence 156444, App
39	70.5	11.2	974	16	US-10-437-963-168282	Sequence 168282, App
40	70.5	11.2	1734	16	US-10-437-963-153570	Sequence 153570, App
41	70	11.1	123	16	US-10-424-599-145465	Sequence 145465, App
42	70	11.1	214	14	US-10-205-194-164	Sequence 164, App
43	70	11.1	417	16	US-10-437-963-122270	Sequence 122270, App
44	70	11.1	518	15	US-10-369-493-3240	Sequence 3240, App
45	70	11.1	518	15	US-10-369-493-3240	Sequence 3240, App

ALIGNMENTS

RESULT 1  
US-10-001-885-125  
Sequence 125, Application US/10001885  
Publication No. US20040058319A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Recipon, Hervé  
APPLICANT: Caffrey, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro  
FILE REFERENCE: DEX-0279  
CURRENT APPLICATION NUMBER: US/10/001,885  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/252,061  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: 60/253,257  
PRIOR FILING DATE: 2000-11-27  
NUMBER OF SEQ ID NOS: 167  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 125  
LENGTH: 121  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-001-885-125  
Query Match 100.0%; Score 632; DB 12; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2,4e-58;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRLIANGELIVQDDPRVRRITTOPPRGSIPIROSFNNRCHGAPPGGPGRQOAGARIGAA 60  
DB 1 MRLIANGELIVQDDPRVRRITTOPPRGSIPIROSFNNRCHGAPPGGPGRQOAGARIGAA 60  
QY 61 OSPFDLRLQVWNGPQWHLGNHAEVPSVILLFILMTGVGGLLVGLVIVSHLSQ 120  
DB 61 OSPFDLRLQVWNGPQWHLGNHAEVPSVILLFILMTGVGGLLVGLVIVSHLSQ 120

QY 121 R 121  
DB 121 R 121

RESULT 2  
US-10-076-747-122  
Sequence 122, Application US/10076747  
Publication No. US20030180725A1  
GENERAL INFORMATION:  
APPLICANT: Salceda, Susana  
APPLICANT: Macina, Roberto  
APPLICANT: Hu, Ping  
APPLICANT: Recipon, Hevra  
APPLICANT: Katta, Kalpana  
APPLICANT: Cafeterkey, Robert  
APPLICANT: Sun, Yongming  
APPLICANT: Liu, Chenghua  
TITLE OF INVENTION: Compositions and Methods Relating to Ovarian Specific Genes and F  
FILE REFERENCE: DEX-0315  
CURRENT FILING DATE: 2002-02-13  
CURRENT APPLICATION NUMBER: 2002-02-13  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 60/266,290  
PRIOR FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 60/266,834  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 122  
LENGTH: 121  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-076-747-122

Query Match 100.0%; Score 632; DB 14; Length 121;  
Best Local Similarity 100.0%; Pred. No. 2,4e-58;  
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYRLANGSEIVODDDPRVTTTTPPPRGSIIPROSFNNGHGAPPGGPPROQOAGARLGAA 60  
DB 1 MYRLANGSEIVODDDPRVTTTTPPPRGSIIPROSFNNGHGAPPGGPPROQOAGARLGAA 60  
QY 61 OSPFDLNRQVLNNGFPGMHGNAVEPVTSIILLFLIMLGVYGLLVGLVYVSHLSQ 120  
DB 61 OSPFDLNRQVLNNGFPGMHGNAVEPVTSIILLFLIMLGVYGLLVGLVYVSHLSQ 120  
QY 121 R 121  
DB 121 R 121

Query Match 61.2%; Score 387; DB 10; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1e-32;

Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYRLANGSEIVODDDPRVTTTTPPPRGSIIPROSFNNGHGAPPGGPPROQOAGARLGAA 60  
DB 39 MYRLANGSEIVODDDPRVTTTTPPPRGSIIPROSFNNGHGAPPGGPPROQOAGARLGAA 98  
QY 61 OSPFDLNRQVLN 73  
DB 99 OSPFDLNRQVLN 111

RESULT 4  
US-09-882-171-673  
Sequence 673, Application US/09882171  
Publication No. US20030175858A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P2  
CURRENT FILING DATE: 2001-05-18  
CURRENT APPLICATION NUMBER: US/09/882,171  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 09/149,476  
PRIOR FILING DATE: 1998-09-08  
PRIOR APPLICATION NUMBER: PCT/US98/04493  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: 60/040,162  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,333  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/038,621  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,626  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,334  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,336  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,163  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/047,600  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,615  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,597  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,502  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,633  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,583  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,617  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,618  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,503  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,592  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,581  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,584  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,500  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,587  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,492  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,598  
PRIOR FILING DATE: 1997-05-23

;; PRIOR APPLICATION NUMBER: 60/047,613  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,582  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,596  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,612  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,632  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,601  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,580  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,568  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,314  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,569  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,311  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,671  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,674  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,669  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,312  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,313  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,672  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,315  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/048,974  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/056,886  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,877  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,889  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,893  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,630  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,878  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,662  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,872  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,882  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,637  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,903  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,888  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,879  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,880  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,894  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,911  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,636  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,874

;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,910  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,864  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,631  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,845  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,892  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/057,761  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/047,595  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,599  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,588  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,585  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,586  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,590  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,594  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,589  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,593  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,614  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,578  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,576  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/047,501  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,670  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/056,632  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,664  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,876  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,881  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,909  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,875  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,862  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,887  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,908  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/048,964  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/057,650  
;; PRIOR FILING DATE: 1997-09-05  
;; PRIOR APPLICATION NUMBER: 60/056,884  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/057,669  
;; PRIOR FILING DATE: 1997-09-05

Query Match 61.2%; Score 387; DB 10; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1e-32;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WRIILANGBIYODDDPRVATTTQPPRGSIPIROS FNRGHBAPGPGPPROQAGARLGA 60

Tue Aug 24 09:48:30 2004

us-10-001-885-125.rapb

Page 4

DB 39 MVRILANGIYODDDPRVTTTOPPRGSIIPROSFNNGHGAAPGPGPPOQAGARLGAA 98  
QY 61 QSPFNDLNRQLYN 73  
DB 99 QSPFNDLNRQLYN 111

RESULT 5  
US-10-164-861-673

Sequence 673, Application US/10164861  
Publication No. US20030225248A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P1  
CURRENT APPLICATION NUMBER: US/10/164,861  
CURRENT FILING DATE: 2002-06-10  
PRIOR APPLICATION NUMBER: US/09/149,476  
PRIOR FILING DATE: 1998-09-08  
PRIOR APPLICATION NUMBER: PCT/US98/04493  
PRIOR FILING DATE: 1998-03-06  
NUMBER OF SEQ ID NOS: 757  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 673  
LENGTH: 111  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-164-861-673

Query Match 61.2%; Score 387; DB 12; Length 111;  
Best Local Similarity 100.0%; Pred. No. 1e-32;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVRILANGIYODDDPRVTTTOPPRGSIIPROSFNNGHGAAPGPGPPOQAGARLGAA 60  
DB 39 MVRILANGIYODDDPRVTTTOPPRGSIIPROSFNNGHGAAPGPGPPOQAGARLGAA 98  
QY 61 QSPFNDLNRQLYN 73  
DB 99 QSPFNDLNRQLYN 111

RESULT 6

US-09-809-391-367  
Sequence 367, Application US/09809391  
Publication No. US20030049618A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P2  
CURRENT APPLICATION NUMBER: US/09/809,391  
CURRENT FILING DATE: 2001-03-16  
PRIOR APPLICATION data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 761  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 367  
LENGTH: 48  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-809-391-367

Query Match 38.8%; Score 245; DB 10; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2.9e-18;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 MGFPQWHLGNHAEVPTISILLFILMLGVRGLLVGLVYVSHLSQR 121  
DB 1 MGFPQWHLGNHAEVPTISILLFILMLGVRGLLVGLVYVSHLSQR 48

RESULT 7

Sequence 367, Application US/09882171  
Publication No. US20030175858A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: 186 Human Secreted proteins  
FILE REFERENCE: P2002P2  
CURRENT APPLICATION NUMBER: US/09/882,171  
CURRENT FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 09/809,391  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 09/149,476  
PRIOR FILING DATE: 1998-09-08  
PRIOR APPLICATION NUMBER: PCT/US98/04493  
PRIOR FILING DATE: 1998-03-06  
PRIOR APPLICATION NUMBER: 60/040,162  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,333  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/038,621  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,626  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,334  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,336  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/040,163  
PRIOR FILING DATE: 1997-03-07  
PRIOR APPLICATION NUMBER: 60/047,600  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,615  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,597  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,502  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,633  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,583  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,617  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,618  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,503  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,592  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,581  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,584  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,500  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,587  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,492  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,598  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,613  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,582  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,596  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,612  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,632  
PRIOR FILING DATE: 1997-05-23  
PRIOR APPLICATION NUMBER: 60/047,601  
PRIOR FILING DATE: 1997-05-23

;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,568  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,314  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,569  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,311  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,671  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,674  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,669  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,312  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,313  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,672  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/043,315  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/048,974  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/056,886  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,877  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,889  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,893  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,630  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,878  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,662  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,872  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,882  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,637  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,903  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,888  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,879  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,880  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,894  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,911  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,636  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,874  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,910  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,864  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,631  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,845  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,892  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/057,761  
;; PRIOR FILING DATE: 1997-08-22

;; PRIOR APPLICATION NUMBER: 60/047,595  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,599  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,588  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,585  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,586  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,590  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,594  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,589  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,593  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/047,614  
;; PRIOR FILING DATE: 1997-05-23  
;; PRIOR APPLICATION NUMBER: 60/043,670  
;; PRIOR FILING DATE: 1997-04-11  
;; PRIOR APPLICATION NUMBER: 60/056,632  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,664  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,876  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,881  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,909  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,875  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,862  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,887  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/056,908  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/048,964  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/057,650  
;; PRIOR FILING DATE: 1997-09-05  
;; PRIOR APPLICATION NUMBER: 60/056,884  
;; PRIOR FILING DATE: 1997-08-22  
;; PRIOR APPLICATION NUMBER: 60/057,669  
;; PRIOR FILING DATE: 1997-09-05

Query Match 38.8%; Score 245; DB 10; Length 48;  
Best Local Similarity 100.0%; Pred. No. 2,9e-18;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 MGPQWHLGNHAEVPTSTILLFLMMLGVRGLLVGLVYVSHLSQR 121  
Db 1 MGPQWHLGNHAEVPTSTILLFLMMLGVRGLLVGLVYVSHLSQR 48

RESULT 8  
US-10-164-861-367  
; Sequence 367, Application US/10164861  
; Publication No. US20030225248A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 186 Human Secreted proteins  
; FILE REFERENCE: P2002P1  
; CURRENT APPLICATION NUMBER: US/10/164,861

Tue Aug 24 09:48:30 2004

us-10-001-885-125.rapb

Page 6

```

; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US/09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 757
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 367
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-164-861-367

Query Match
Best Local Similarity 38.8%; Score 245; DB 12; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.9e-18;
Matches 48; Conservative 0; Mismatches 0; Gaps 0;
```

```

QY 74 MGFPQMHGNHAYEPTSTILLFLMLGVRLGVLYVSHLSQR 121
DB 1 MGFPQMHGNHAYEPTSTILLFLMLGVRLGVLYVSHLSQR 48
```

```

RESULT 9
US-10-108-260A-3882
; Sequence 3882, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3882
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3882
```

```

Query Match
Best Local Similarity 17.5%; Score 110.5; DB 15; Length 132;
Best Local Similarity 31.9%; Pred. No. 0.0012;
Matches 30; Conservative 17; Mismatches 34; Indels 13; Gaps 3;

QY 26 RGSIPRQS-----FNRGHGAPPGGPRQQAAGALGAASPFNDLNROLVNMGFPQMTL 81
DB 42 RGRPKSEQVEJESQNTGEPVGDYK-----KMGTL--FGLKNKLNMGFTMYF 92
QY 82 GNAHVEPTSTILLFLMLGVRLGVLYVSHLSQR 115
DB 93 GERIVEPTVIFFWVLMFLGLQALGLAVAVLCV 126
```

```

RESULT 10
US-10-156-761-10980
; Sequence 10980, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HOSHIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
```

```

; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10980
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Streptomyces avermiltillis
US-10-156-761-10980
```

```

Query Match
Best Local Similarity 15.7%; Score 99.5; DB 14; Length 223;
Best Local Similarity 30.9%; Pred. No. 0.031; 21; Indels 35; Gaps 4;
Matches 29; Conservative 9; Mismatches 21;

QY 23 QPP-----RGSIPRQS-----FNRGHGAP--PGGPRQQAAGALGAASPFNDLNROL 71
DB 27 QPPYQAPGGRIPQGGYXGYPQGGPGYXGYPGGPGGPRVAMGRLGA----- 74
QY 72 VNMGFPQMHGNHAYEPTSTILLFLMLGVRLGV 105
DB 75 -----RAIDGVAFFVTFPLGALGVAG 96
```

```

RESULT 11
US-10-424-599-174474
; Sequence 174474, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 174474
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_128569C.1.pep
US-10-424-599-174474
```

```

Query Match
Best Local Similarity 13.5%; Score 85.5; DB 12; Length 112;
Best Local Similarity 32.8%; Pred. No. 0.4;
Matches 22; Conservative 8; Mismatches 26; Indels 11; Gaps 2;

QY 37 RGHGAPPGGPRQQAAG--ARLGAASPFND-----LNROLVNMGFPQMHGNH 85
DB 5 RRGSGPPGKGRSAAKGRAGAKPPNPULVFPAGVILLNKQPPFALAPMKIPIVSK 64
QY 86 VEPVTSI 92
DB 65 SEPIAI 71
```

```

RESULT 12
US-10-437-963-198292
; Sequence 198292, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
```



```

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 198292
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_93967C.1.pcp
US-10-437-963-198292

Query Match      13.0%; Score 82; DB 16; Length 123;
Best Local Similarity 42.9%; Pred. No. 1;
Matches 27; Conservative 4; Mismatches 14; Indels 18; Gaps 5;

QY 17 RVRTTTPPGSIFRSGF-FNRHGAP-----PGGPGPQQGAGAR----LGAA 60
DB 25 RGRTRT-PTRGSPRRSAFPFRGPGPPRQGLRAGAGGP-PPGRRGARGAPLHVAG 82

QY 61 QSP 63
DB 83 KSP 85

RESULT 13
US-10-425-114-60982
; Sequence 60982, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60982
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3356-039-A12_F11.pcp
US-10-425-114-60982

Query Match      12.8%; Score 81; DB 12; Length 256;
Best Local Similarity 33.0%; Pred. No. 3.1;
Matches 29; Conservative 7; Mismatches 26; Indels 26; Gaps 4;

QY 26 RGSIPRGSFNRHGAPPGGPGP-----RQQAARIGA-----AQSFPNDLNR 69
DB 81 RGAUFR--LIRGGAGQGVPPDARRRGLRRRRRRRAGARGAHPRAQVHGRRHPLRLH 137

QY 70 QLVNM-----GPPQHLGNHAEVPT 90
DB 138 RHVRRALRHENGDPPDERGVHAAPVT 165

RESULT 14
US-10-106-698-8371
; Sequence 8371, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
```

```

; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 8371
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (88)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (89)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (91)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-8371

Query Match      12.7%; Score 80; DB 14; Length 92;
Best Local Similarity 35.7%; Pred. No. 1.2;
Matches 30; Conservative 2; Mismatches 26; Indels 26; Gaps 4;

QY 15 DPRVR-----TTTPRGSIFRSGF-FNRHGAPPGGPGPQQGAGARGAASFPND 66
DB 15 DPRVRRSSGSPLYTRVFRAS-PAPEGPSGHALPGAGLR--AGHERGAAR---- 66

QY 67 LNRQLVNMGFPQWHLGNHAEVPT 90
DB 67 -----PRAHGSRSAARRGT 80

RESULT 15
US-10-156-761-10160
; Sequence 10160, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OKURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10160
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10160

Query Match      12.7%; Score 80; DB 14; Length 392;
Best Local Similarity 27.2%; Pred. No. 6.6;
Matches 37; Conservative 18; Mismatches 33; Indels 48; Gaps 8;

QY 7 NGEIVQDDDPRT-----TTTPRGSIFRSGF-FNRHG-----APPGGPGP 48
```

Tue Aug 24 09:48:30 2004

us-10-001-885-125.rapb

Page 8

```
Db      67  SGPLFRDDTPQTTSYGSQTPYVTOSPYGS-----QTPHGAOPQKPEPMDAPQAPAP 119
Oy      49  ROOQAGARLGAAGSPFNDLNRQLVNMGFPOWHLGNHAEVTSILLFL--MMLGVRL 106
Db      120  QKSNAGRDLAGAIG-----VGVG-----LG-----VVIYASLFFVKAFVGVIAV 159
Oy      107  -LIVGLVYLVSHLSOR 121
Db      160  AVVVGILWELTSRIER 175
```

Search completed: August 24, 2004, 09:33:21  
Job time : 126 secs

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2004, 07:11:30 ; Search time 2048 Seconds

(without alignments)  
2560.794 Million cell updates/sec

Title: US-10-001-885-125

Sequence: 1 MVTILNGELTVQDDPRVRT.....GVKGLLVGVLYVLSHSQR 121

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xld  
-O=/cgrn1/USPFO.spool/p/US10001885/runat\_17082004\_151718\_25904/app\_query\_faeta\_1.263  
-DB=genmbd1 -QWMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10001885 @CGN 1.1 5600 @runat\_17082004\_151718\_25904 -NCPU=6 -ICPU=3  
-NO MAP -LARGEOUTRY -NEG\_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAXN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Genmbd1:\*  
1: gb ba:\*  
2: gb hcg:\*  
3: gb in:\*  
4: gb om:\*  
5: gb ov:\*  
6: gb pat:\*  
7: gb ph:\*  
8: gb pl:\*  
9: gb pr:\*  
10: gb ro:\*  
11: gb sts:\*  
12: gb sy:\*  
13: gb un:\*  
14: gb vl:\*  
15: em ba:\*  
16: em fun:\*  
17: em hum:\*  
18: em in:\*  
19: em mu:\*  
20: em om:\*  
21: em ov:\*  
22: em or:\*  
23: em pat:\*  
24: em ph:\*  
25: em pl:\*  
26: em ro:\*  
27: em sts:\*  
28: em un:\*

29: em vl:\*  
30: em hcg hum:\*  
31: em hcg inv:\*  
32: em hcg other:\*  
33: em hcg mus:\*  
34: em hcg pln:\*  
35: em hcg rod:\*  
36: em hcg mam:\*  
37: em hcg vrt:\*  
38: em sy:\*  
39: em hcg hum:\*  
40: em hcg mus:\*  
41: em hcg other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	632	100.0	1049	6 BD191093	BD191093 186 human
2	632	100.0	1129	6 BC013587	BC013587 Homo sapi
3	632	100.0	1192	6 AX431088	AX431088 Sequence
4	577	91.3	550	6 AX113082	AX113082 Sequence
5	568.5	90.0	1538	10 BC024943	BC024943 Mus muscu
6	563.5	89.2	275924	2 AC116233	AC116233 Rattus no
7	551	87.2	483	6 AR413227	AR413227 Sequence
8	551	87.2	483	6 BD108780	BD108780 EST and e
9	544.5	86.2	240657	2 AC111675	AC111675 Rattus no
10	503	79.6	519	6 AR415427	AR415427 Sequence
11	503	79.6	519	6 BD110980	BD110980 EST and e
12	472	74.7	165110	9 AL450311	AL450311 Human DNA
13	472	74.7	173341	2 AC021954	AC021954 Homo sapi
14	459	72.6	215050	2 AC127417	AC127417 Mus muscu
15	448	70.9	258815	2 AC127817	AC127817 Rattus no
16	443.5	70.2	138070	2 AC109783	AC109783 Mus muscu
17	135	21.4	754	3 AK113891	AK113891 Clona int
18	130	20.6	238937	2 AC094926	AC094926 Rattus no
19	110.5	17.5	2594	6 AX834315	AX834315 Sequence
20	110.5	17.5	2594	6 AK096689	AK096689 Homo sapi
21	110	17.4	2565	9 BC022534	BC022534 Homo sapi
22	107.5	17.0	109668	2 AC109347	AC109347 Homo sapi
23	107.5	17.0	170662	2 BX005193	BX005193 Danto rer
24	107.5	17.0	172287	2 AC069028	AC069028 Homo sapi
25	107	16.9	155320	2 AC135762	AC135762 Rattus no
26	107	16.9	241394	2 AC131549	AC131549 Rattus no
27	105	16.6	496	10 BC039568	BC039568 Mus muscu
28	105	16.6	151241	2 AC115041	AC115041 Mus muscu
29	105	16.6	153126	2 AC102465	AC102465 Mus muscu
30	101	16.0	1383	6 AR385066	AR385066 Sequence
31	101	16.0	212840	2 AC135223	AC135223 Gallus ga
32	99.5	15.7	183599	2 AC073501	AC073501 Homo sapi
33	99.5	15.7	296300	1 AP005035	AP005035 Streptomy
34	98.5	15.6	149610	8 AP003244	AP003244 Oryza sat
35	98.5	15.6	163513	8 AP003200	AP003200 Oryza sat
36	96.5	15.3	66254	9 AL137791	AL137791 Human DNA
37	96.5	15.3	179321	9 AC114493	AC114493 Homo sapi
38	95	15.0	3705	5 GGA53435	GGA53435 Gallus ga
39	94.5	15.0	10348	1 AB004846	AB004846 Pseudomon
40	94.5	15.0	177226	1 AC121858	AC121858 Mus muscu
41	94.5	15.0	183975	10 AL590988	AL590988 Mouse DNA
42	94.5	15.0	248358	2 AC122944	AC122944 Rattus no
43	93.5	14.8	156997	9 AC011465	AC011465 Homo sapi
44	93	14.7	236639	2 AC127406	AC127406 Rattus no
45	93	14.7	254489	2 AC098275	AC098275 Rattus no

RESULT 1

ALIGNMENTS

BD191093

LOCUS

BD191093 1049 bp DNA linear PAT 17-JUL-2003

DEFINITION

186 human secreted proteins.

ACCESSION

BD191093.1 GI:33000832

VERSION

JP 2002510192-A/57.

KEYWORDS

unidentified

SOURCE

unclassified

REFERENCE

AUTHORS

1 (bases 1 to 1049)  
Ruben,S.M., Rosen,C.A., Fischer,C.L., Soppet,D.R., Carter,K.C.,  
Bednarik,D.P., Endress,G.A., Yu,G.L., N.J., Feng,P., Young,P.E.,  
Greene,J.M., Ferris,A.M., Duan,R., Hu,J.S., Florence,K.A.,  
Olson,H.S., Bbner,R., Brewer,L.A., Moore,P.A., Shi,Y.,  
Lafleur,D.W., Li,Y., Zeng,Z. and Kyaw,H.

COMMENT

Patent: JP 2002510192-A 57 02-APR-2002;  
HUMAN GENOME SCIENCES INC  
JP 2002510192-A/57

TITLE

186 human secreted proteins  
Patent: JP 2002510192-A 57 02-APR-2002;  
HUMAN GENOME SCIENCES INC  
JP 2002510192-A/57

JOURNAL

02-APR-2002  
06-MAR-1998 JP 1998538883  
07-MAR-1997 US 60/040162, 07-MAR-1997 US 60/040333 PR  
07-MAR-1997 US 60/038621, 07-MAR-1997 US 60/040161 PR  
07-MAR-1997 US 60/040626, 07-MAR-1997 US 60/040334 PR  
07-MAR-1997 US 60/040336, 07-MAR-1997 US 60/040163 PR  
11-APR-1997 US 60/043580, 11-APR-1997 US 60/043568 PI  
M RUBEN, CRAIG A ROSEN, CARIE L FISCHER, DANIEL R SOPPET, PI  
KENNETH C CARTER, DANIEL P BEDNARIK, GREGORY  
A ENDRESS, GUO LIANG  
PI YU JIAN NI  
PI PING FENG, PAUL E YOUNG, JOHN M GREENE, ANN  
M FERRIS, ROXANNE DUAN,  
PI JING SHAN HU, KIMBERLY A FLORENCE, HENRIK  
S OLSEN, REINHARD EBNER,  
PI LAURIE A BREWER, PAUL A MOORE, YANGGU SHI, DAVID W LAFLEUR PI  
YI LI ZHIZHEN ZENG,  
PI HUA KYAW  
PC C12N15/12, C12N5/10, C12N1/21, C07K14/47, C07K16/18, C12Q1/68, PC  
G01N33/50,  
PC G01N33/53, G01N33/68, A61K38/17  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.

FEATURES

Source Location/Qualifiers.  
1..1049  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

ORIGIN

Alignment Scores: 1,52e-43 Length: 1049  
Pred. No.: 632.00 Matches: 121  
Score: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-001-885-125 (1-121) x BD191093 (1-1049)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 116 ATGTGGCGGATCTTGGCCAAATGGGGAATCGTGCAGCAGCAGCCCCCGAGTGAACACC 175  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyVH:Gly 40  
DB 176 ACTACCCAGCCACCAAGAGTGAATCTTCGACAGACCTTCTTCATATGAGGAGCAAGGT 235  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 236 GTTCCCCAGAGGGGCTCTGGCCCCCGCCAGCAGCAGGAGGTGCCAGGCTGGTGTGCT 295

REMARK

COMMENT  
NIG-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. W.

Clone distribution: MGC clone distribution information can be found  
through the I M A G E Consortium/ILNL at: <http://image.llnl.gov>

DB

296 CAGTCCCTTCAAGACCTCAACCGCAGCTGGGAAACAGGGGTTCCGACGAGCAT 355

QY

81 LeuGlyAamHhAlaValGluProValThrSerIleLeuLeuPheLeuMetMet 100

DB

356 CTGGGACATGCTCTGTGAGCGGTGACCTTCATCTCTCTCTCTCTCATATG 415

QY

101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValIYLeuValSerHisLeuSerGln 120

DB

416 CTTGGTGTCTGGGCTCTCTCTGTGGCTTGTCTACCTGAGTGTCCACCTGATGTCAG 475

QY

121 Arg 121

DB

476 CGG 478

RESULT 2

BC013587 1129 bp mRNA linear PRI 11-DEC-2003  
LOCUS  
DEFINITION  
Homo sapiens chromosome 10 open reading frame 35, mRNA (cdna clone  
MGC:9596 IMAGE:389656), complete cds.

ACCESSION

BC013587

VERSION

BC013587.1 GI:15488919

KEYWORDS

MGC.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

1 (bases 1 to 1129)  
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,K., Marusina,K., Farmer,A.A., Casavant,T.L.,  
Stapleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L.,  
Scheetz,T.E., Brownstein,K.J., Ubedin,T.B., Toshtylyk,S.,  
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
Abramson,R.D., Mallaby,S.J., Bosak,S.A., McSwan,P.J.,  
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Holly,S.M.,  
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
Raney,V., Helton,E., Kerteman,M., Madan,A., Rodriguez,S.,  
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
Boutfard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
Butterfield,Y.S., Krzywinski,M.I., Skalka,J., Smalls,D.B.,  
Schnerb,A., Schein,J.B., Jones,S.J. and Marra,M.A.

TITLE

Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL

Strausberg,R.

PUBMED

12477932

REFERENCE

2 (bases 1 to 1129)

AUTHORS

Direct Submission

TITLE

Submitted (04-SEP-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

JOURNAL

NIG-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. W.



misc\_feature /mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
1..550  
/note="n = A,T,C or G"

## ORIGIN

Alignment Scores:  
Pred. No.: 2,76e-39 Length: 550  
Score: 577.00 Matches: 117  
Percent Similarity: 97.52% Conservative: 1  
Best Local Similarity: 96.69% Mismatches: 2  
Query Match: 91.30% Indels: 2  
DB: Gaps: 0

US-10-001-885-125 (1-121) x AX113082 (1-550)

CY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 115 ATGTGGCGGATCTTGGCCCATGGGAAATCTGCAGACACACACCCCGAGAGACC 174  
CY 21 ThrThrgInProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 175 ACTACCCAGCAGCAGCAGGAGTAGCATTCCTCGACAGAGCTTCTCATAGGGCCCATGTGT 234  
CY 41 AAlaProProGlyGlyProGlyProArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60  
DB 235 GCTCCCCCGGGGGGCTCTGGCCCCCGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 294  
CY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnITPHis 80  
DB 295 CAGTCCCCCTTCATGACCTCAACCGGAGCTGTGAACATGGGCTTCCGCAAGTGGCAT 354  
CY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuLeuPheLeuMetMet 100  
DB 355 CTGGGCAACCATGCTGTGGAGCCGGTACCTCATCTCTGTCTTCTGCTCATGAT 413  
CY 100 tLeuGlyValArgGlyLeuLeuValGlyLeuValGlyLeuValSerHisLeuSerG 120  
DB 414 GCTTGCTGCTGCTGCTGCTTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 473  
CY 120 n 120  
DB 474 A 474

RESULT 5  
BC024943  
LOCUS  
DEFINITION Mus musculus RIKEN cDNA 2010107G23 gene, mRNA (CDNA clone MGC:28201  
IMAGE:3989515), complete cds.  
ACCESSION BC024943  
VERSION BC024943.1 GI:19354288  
KEYWORDS MGC.  
SOURCE Mus musculus (house mouse)  
ORGANISM  
REFERENCE  
AUTHORS

Strusberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Nak S.I., Wang J., Heide F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stableton M., Soares M.B., Bonaldo M.F., Casavant T.L.,  
Schetz T.E., Brownstein M.J., Usdin T.B., Toshitsugu S.,  
Carroll P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,  
Abramsen R.D., Mullaly S.J., Bosak S.A., McEwan P.J.,  
McKernan K.J., Malek J.A., Gamarale P.H., Richards S.,  
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S.,  
Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,  
Bonfield G., Blakeley R.W., Touchman J.W., Green B.D.,

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

## REMARK

## COMMENT

Butterfield, Y.S., Krzywinaki, M.I., Skalska, U., Smalins, D.E.,  
Schmerch, A., Schein, J.E., Jones, S.J., and Marr, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
2 (bases 1 to 1538)  
Strusberg, R.  
Direct Submission  
Submitted (01-MAR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMT)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louis, H.,  
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nambavi,  
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILMT at: <http://image.llnl.gov>  
Series: IRAX Plate: 36 Row: h Column: 13  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Similarity but not  
identity to protein.

## FEATURES

## source

1..1538  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="MGC:28201 IMAGE:3989515"  
/tissue="mammary tumor metastasized to lung. Tumor  
arose spontaneously from a senescent normal mammary  
(clonal) outgrowth infected with the virus MMTV."  
/clone\_1lb="NCI CGAP\_Lu29"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..1538  
/gene="2010107G23R1k protein"  
/db\_xref="LOCUSID:69894"  
/db\_xref="MGI:1917144"  
486..848  
/codon\_start=1  
/product="2010107G23R1k protein"  
/protein\_id="AAH24943.1"  
/db\_xref="GI:19354289"  
/db\_xref="LOCUSID:69894"  
/translation="MVRILANGELIVQDDPRVYRTTQGRSSSGGFFRRRGHGAAPGG  
PGRPOOAGARLGAAGSPSILNRQLVNMGPQWHLGNHVPVTSILLPLFMKVLGV  
RGLLVGLVTVSHLSQR"

## CDS

ORIGIN  
Alignment Scores:  
Pred. No.: 4.08e-38 Length: 1538  
Score: 568.50 Matches: 112  
Percent Similarity: 94.21% Conservative: 2  
Best Local Similarity: 92.56% Mismatches: 6  
Query Match: 89.95% Indels: 1  
DB: Gaps: 1

US-10-001-885-125 (1-121) x BC024943 (1-1538)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnSerPheAspProArgValArgThr 20  
 Db 486 ATGGAGAGATCTGGCCCAATGGGGAGATCTTCAAGATATGACCCACAGATGAGACG 545  
 QY 21 ThrThrglnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40  
 Db 546 ACCACCCAG---CAAGAGAGTAGAGCTCTCAGAGGGCTTTTTCACAGAGGGCCAGCGC 602  
 QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaGlyLeuGlyAla 60  
 Db 603 GCACCTCCAGGGGGCCCTGGACCCCGCAGCAGCAGGAGTGTCCCACTGGGTGCTGCC 662  
 QY 61 GlnSerProPheAsnAspLeuAsnArgGlyIleuValAsnMetGlyPheProGlnTrpHis 80  
 Db 663 CAATCTCTTTCAGTGCACCTGACCGGACGCTGCTGCAATGGGCTTCCCAATGGCAGC 722  
 QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
 Db 723 CTGGAGAACCGCGTGGAGACCTGTAACCTTCACTCTCTCTCTCTCTCTCTATATG 782  
 QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrluValSerHisLeuSerGln 120  
 Db 783 CTCGGGGGTTCTGGCT 842  
 QY 121 Arg 121.  
 Db 843 CGG 845  
 RESULT 6  
 AC116233/275924 bp DNA linear HTG 22-SEP-2002  
 LOCUS Rattus norvegicus clone CH230-95L17, \*\*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\*\*, 4 unordered pieces.  
 ACCESSION AC116233  
 VERSION AC116233.7 GI:23101097  
 HTG: HTGS PHASE1; HTGS DRAFT; HTGS ENRICHED.  
 KEYWORDS Rattus norvegicus (Norway rat)  
 SOURCE Rattus norvegicus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 REFERENCE 1 (bases 1 to 275924)  
 Muzny,D.,Marler,M.,Lee,A.,Abramson,S.,Adams,C.,Alder,J.,Allen,C.,Allen,H.,Alshrocks,S.,Amin,A.,Anguiano,D.,Aryalabechi,V.,Ayagi,A.,Ayodeji,M.,Baba,E.,Baden,H.,Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,Biswal,N.,Blair,D.,Blankenburg,K.,Blaych,P.,Brown,M.,Bryant,N.,Buhay,C.,Burck,P.,Burrell,K.,Calderon,E.,Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Center,A.,Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,Davis,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dederich,D.,Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Divya,K.,Diaper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Dvali,B.,Eaves,K.,Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,Fraser,C.M.,Gabisi,A.,Ganta,R.,Garcia,A.,Garnier,T.,Gavara,M.,Gebregeorgis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,M.,Guevara,M.,Guhartne,P.,Haaland,M.,Hamil,C.,Hamilton,C.,Hamilton,K.,Harvey,Y.,Havlik,P.,Hawes,A.,Henderson,N.,Hernandez,J.,Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,Hollins,B.,Howells,S.,Huliyil,S.,Hume,J.,Idelberg,D.,Jackson,A.,Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolivet,A.,Karpach,S.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Kovar,C.,Kraus,C.,Kraft,C.L.,Lebow,H.,Levan,J.,Lewis,L.,Li,Z.,Liu,J.,Liu,Y.,Liu,M.,Liu,Y.,London,P.,Longacre,S.,Lopez,D.,Lorenschwartz,M.,Mahindaratne,M.,Mahmoud,M.,Malloy,K.,Mangum,A.,Mangum,B.,Mapua,P.,Martin,K.,Martin,R.,Martinez,E.,Manguney,S.,McLeod,M.P.,McNeill,T.Z.,Meenan,E.,Milosavljevic,A.,Miner,G.,Minja,E.,Montemayor,J.,Moore,S.,Morgan,M.,Morris,K.,Morris,S.,Munidas,M.,Murphy,M.,Nairi,L.,Nankervis,C.,Neal,D.,Newton,N.,Nguyen,N.,Norris,S.,

Nwackeme, O., Okwunna, G., Olarunbasogun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Polindexter, A., Popovic, D., Pimm, E., Pu, L., L., Puzo, M., Quito, J., Rachlin, B., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ritz, S., J., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Wortley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhuesern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 275924)  
 Morley, K.C.  
 Direct Submission  
 Submitted (26-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 275924)  
 Rat Genome Sequencing Consortium.  
 Submitted (22-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On Sep 18, 2002 this sequence version replaced gi:21738715.  
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GRF  
 Center clone name: CH230-95L17  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 234606 bases at least Q40  
 Consensus quality: 237265 bases at least Q30  
 Consensus quality: 238894 bases at least Q20  
 Estimated insert size: 264374; sum-of-contigs estimation  
 Quality coverage: 3x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html)  
 \* NOTE: This sequence may represent more than one clone.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1  
 \* 10838: contig of 10838 bp in length  
 \* 10839 10938: gap of unknown length  
 \* 10939 273385: contig of 263347 bp in length  
 \* 273386 273385: gap of unknown length  
 \* 273386 274618: contig of 1233 bp in length  
 \* 274619 274718: gap of unknown length

FEATURES	* 274719	275924: contig of 1206 bp in length.
Location/Qualifiers	1..275924	
source	/organism="Rattus norvegicus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10116"	
	/clone="CH230-95L17"	
misc_feature	2610..3635	
	/note="wgs_end_extension"	
misc_feature	clone_end:Sp6"	
	/note=".10838"	
misc_feature	/note="wgs_end_extension"	
	clone_end:Sp6"	
	10939..12161	
	/note="wgs_end_extension"	
	clone_end:Sp6"	
misc_feature	91900..93203	
	/note="wgs_end_extension"	
	clone_end:Sp6"	
	135055..135316	
	/note="clone boundary"	
	clone_end:Sp6"	
	site:EcORI	
	end_sequence:BH361167"	
misc_feature	165585..210774	
	/note="clone boundary"	
	clone_end:T7	
	site:EcORI	
	end_sequence:BH361166"	
ALIGNMENT SCORES:		
Pred. No.:	2,386-35	Length: 275924
Score:	563.50	Matches: 111
Percent Similarity:	93.39%	Conservative: 2
Best local Similarity:	91.74%	Mismatches: 7
Query Match:	89.16%	Indels: 1
DB:	2	Gaps: 1
US-10-001-885-125 (1-121) x AC116233 (1-275924)		
Qy	1 MetValArglleuulaasnlgylulilevalglnaapaaapProArgValArgThr	20
Db	227956 ATGGTGAAGATCTGGCTAAATGGGAGATCGTTCAAGATGATGACCCCGAGTGAAGACT	227897
Qy	21 ThThrGlnProArgArglySerllePoaArgGlnSerPhePheAsnArgGlyHnlsly	40
Db	227896 ACCACCCCA---CAGGAGAGTGAAGACACCCGCGAGGCTTTTTCACAGAGGCACAGT	227840
Qy	41 AlaProProGlylylProGlyProArgGlnGlnGlnalaglyalaArgLeuGlyalaala	60
Db	227839 GCACCTCCAGGGGGCCCTGTGAACCCCGCAGCAGCAGGAGGTGCCGACGTGGGCTGCC	227780
Qy	61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValaMetGlyPheProGlnITPRHis	80
Db	227779 CAATCTCCCTTTCAGTCACTGAAACCGGACGCTGTGAACATGGGCTTCCCACAATGGCAT	227720
Qy	81 LeuGlyAsnHisAlaValGlnProValThrSerlleLeuLeuPheLeuMetMet	100
Db	227719 CTTGGGAACACGCTGTGAAGCTGTGAACCTCATCTCCGTCTTCTGTCTCATGATA	227660
Qy	101 LeuGlyValArglyleuLeuLeuValGlyLeuValIlyLeuValSerHisLeuSerGln	120
Db	227659 CTGGGGGCTTGTGGCTCTCGCTTGGGTCTGTGCTCACTGGTGTCTCACTTGAGCCAG	2276000
Qy	121 Arg 121	
Db	227599 CGG 227597	
RESULT 7		
AR413227	483 bp	DNA linear PAT 18-DEC-2003
LOCUS		

ACCESSION	AR413227
VERSION	AR413227.1 GI:40168337
KEYWORDS	
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 483)
TITLE	Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
JOURNAL	EST's and encoded human proteins
FEATURES	Patent: US 6639063-A 864 28-OCT-2003;
Source	Location/Qualifiers
	1..483
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
Alignment Scores:	
Pid. No.:	3,42e-37 Length: 483
Score:	551.00 Matches: 104
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	87.18% Indels: 0
DB:	Gaps: 0
US-10-001-885-125 (1-122) x AR413227 (1-483)	
Oy	1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20
Dd	171 ATGTGGCGAATCTTGGCCAAAGGGGAATCGTGAGAAGCAGACCCTTCATTGAGGCCATGC 230
Oy	21 ThrThrgInPProArGLysErILePrOArGLInserPheAsnArGLyHISgly 40
Dd	231 ACTACCAAGCACCAAGAGTAGCATTTCTCGACGAGCTTCTTCATAATAGGGCCATGCT 290
Oy	41 AlaProPProGIgLYIPrOGLyIPrOArGLIngInglalaglyalaargleuglyvalala 60
Dd	291 GCTCCCCCAGGGGGGTCTGGCCCCCGCAGAGAGAGAGGTGCGAGCTGGGTGCT 350
Oy	61 GLInSerPProPHeAsnAPLeuASnArgGLInLeuValasMetGLYPnePROGLITrpHis 80
Dd	351 CAGTCCCCTTCATGACTCAACCGGAGCGTGTGAACAATGGGCTTCCGCAATGGCAT 410
Oy	81 LeuGLyaenhiSaLaVaLgLnProvAlThisErILEuleuLeuPheLeuMetNet 100
Dd	411 CTCGGCAACCACTGTGTGGAGCCGCTGACCTCCACCTCCCTCTCTCTCATGATG 470
Oy	101 LeuGIyValArg 104
Dd	471 CTGGTGTTCCGT 482
RESULT 8	
LOCUS	BD108780 483 bp DNA linear PAT 18-SEP-2002
DEFINITION	EST and encoded human protein.
ACCESSION	BD108780 BD108780.1 GI:23203598
VERSION	UP 2002010789-A/857.
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 483) Edwards,J.B.D.M., Jobert,S. and Giordano,J.E. EST and encoded human protein Patent: JP 2002010789-A 857 15-JAN-2002;
AUTHORS	
TITLE	
JOURNAL	GENSET CORP
COMMENT	OS Homo sapiens (human) PN UP 2002010789-A/857 PD 15-JAN-2002 PF 07-AUG-2000 JP 2000280989 PR 05-AUG-1999 US 60/147499 PI JEAN BAPTIST DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI



PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N15/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC  
C12N15/00  
CC EST and encoded human protein  
FH Key Location/Qualifiers  
FT CDS 171..482.  
1.483  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ALIGNMENT SCORES:  
Pred. No.: 3,42e-37 Length: 483  
Score: 551.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 87.18% Indels: 0  
Gaps: 0

US-10-001-885-125 (1-121) x BD108780 (1-483)  
QY 1 MetValArgIleLeuAlaAsnGlyValLeuValGlnAspAspArgProArgValArgThr 20  
DB 171 ATGGGCGGATCTTGCCCAATGGGAAATCTGTGACGACGACGACCCCGGATGAGGACC 230  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyValIleGly 40  
DB 221 ACTACCCGACCAACCAAGAGGTAGCATCTCTCGACAGAGCTTCTCAATAGGGCCCATGCT 290  
QY 41 AlaProProGlyValProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 291 GCTCCCGACGAGGGGCTCTGGCCCGCCGACGACGAGGAGGTGCGAGGCTGGGCTGCT 350  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAlaMetGlyPheProGlnTrpHis 80  
DB 351 CAGTCCCTTCAATGACCTCAACCGGACGCTGTGGAACATGGGCTTCCGACATGGCAT 410  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuLeuPheLeuMetMet 100  
DB 411 CTCGGCAACCATGCTGTGGAGCCGCTGACCTTCATCTCTCTCTCTCTCATGATG 470  
QY 101 LeuGlyValArg 104  
DB 471 CTGTGTCTCT 482

RESULT 9  
AC111675/c 240657 bp DNA linear HTG 13-MAY-2003  
LOCUS Rattus norvegicus clone CH230-48G14, WORKING DRAFT SEQUENCE, 4  
DEFINITION unorderd pieces.  
ACCESSION AC111675.5 GI:30579217  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULFILL.  
KEYWORDS Rattus norvegicus (Norway rat)  
SOURCE Rattus norvegicus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 240657)  
Muzny, D., Maric, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, R.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Dalla, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K.,

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gerregeorgis, E., Geier, K., Giller, R., Grady, M., Guerra, W., Guvata, W.,  
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladik, S.L., Hodgson, A., Hogues, M.,  
Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, H., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C.L., Ledow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, D., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorenz, J., Louie, H., Lozano, R.J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenan, E.,  
Milosavljevic, A., Miner, G., Minig, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Muniasa, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwankwelu, O., Okunolu, G., Olapadesogun, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,  
Pioppert, F., Polindexter, A., Popovic, D., Prims, E., Pu, L.,  
Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Register, M.A., Reigh, R.,  
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,  
Sanders, W., Savary, G., Scherer, S., Scott, G., Shattman, S., Shen, H.,  
Shetty, V., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smales, D.,  
Snead, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
Stemle, M., Strong, R., Sutton, A., Svarek, A., Tabor, P., Taylor, C.,  
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J.,  
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,  
Williams, G., Willson, R., Wleczky, R., Woodson, H., Woreley, K.,  
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhou, S., Dunn, D., von  
Weinstock, G., and Gibbs, R.A.  
Direct Submission  
Unpublished  
2 (bases 1 to 240657)  
Worley, K.C.  
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 240657)  
Rat Genome Sequencing Consortium.  
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On May 13, 2003 this sequence version replaced gi:24818892.  
The sequence in this assembly is a combination of BAC based reads  
and whole genome shotgun sequencing reads assembled using Atlas  
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
in the feature table below represents a scaffold in the Atlas  
assembly (a 'contig-scaffold'). Within each contig-scaffold,  
individual sequence contigs are ordered and oriented, and separated  
by sized gaps filled with Ns to the estimated size. The sequence  
may extend beyond the ends of the clone and there may be sequence  
contigs within a contig-scaffold that consist entirely of whole  
genome shotgun sequence reads. Both end sequences and whole genome  
shotgun sequence only contigs will be indicated in the feature  
table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GWT  
Center clone name: CH230-48G14

----- Summary Statistics -----  
Assembly program: Atlas 3.0;  
Consensus quality: 228778 bases at least Q40  
Consensus quality: 231963 bases at least Q30  
Consensus quality: 234139 bases at least Q20  
Recloned insert size: 237555; sum-of-coverage estimation  
Quality coverage: 7x in Q20 bases; sum-of-coverage estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 236811: contig of 236811 bp in length  
\* 236812 236811: gap of unknown length  
\* 236912 238255: contig of 1344 bp in length  
\* 238256 238355: gap of unknown length  
\* 238356 239505: contig of 1150 bp in length  
\* 239506 239605: gap of unknown length  
\* 239606 240657: contig of 1052 bp in length.

## FEATURES

source

1.240657  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-48G14"  
1..2145

misc\_feature

/note="wgs\_end\_extension  
clone\_end:Sp6"

misc\_feature

/note="wgs\_end\_extension  
clone\_end:Sp6"

misc\_feature

8292..9121  
/note="clone boundary  
clone\_end:Sp6"

misc\_feature

end\_sequence:BH363875"  
/note="wgs\_end\_extension  
clone\_end:T7"

misc\_feature

end\_sequence:BH363875"  
/note="wgs\_end\_extension  
clone\_end:T7"

## ORIGIN

Alignment Scores:

Pred. No.: 7 71e-34 Length: 240657  
Score: 544.50 Matches: 110  
Percent Similarity: 91.80% Conservative: 2  
Best Local Similarity: 90.16% Mismatches: 8  
Query Match: 86.16% Indels: 2  
DB: 2 Gaps: 1

US-10-001-885-125 (1-121) x AC111675 (1-240657)

QY 1 MetValArgIleuValAsnGlyGluIleValGlnAspAspProArgValArgThr 20

Db 51622 ATGGTACGATCTGGCTATGAGGAGATCGTTCAGATGATGACCCCGAGTGGAGCT 51663

QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40

Db 51562 ACCACCCAA---CAGAGAGTACGACACCCGACGGCTTTTCAACAGAGCCACAGT 51596

QY 41 AlaProProGlyIlePro-GlyProArgGlnGlnAlaGlyAlaGlyLeuGlyAlaI 60

QY 60 agInserProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHi 80

Db 51445 CCAATCTCTTCACTGACCTGACACCGGACCTGTGAACATGGCTTCCACATGGCA 51386

QY 80 sLeuGlyAsnHisAlaValGluProValTrpSerIleLeuLeuPheLeuMetMet 100

Db 51385 TCTTGGGAACACGATGAGGCTGTGACCTCCATCTCTCTCTCTCTCATGAT 51326

QY 100 tLeuGlyValArgGlyLeuLeuLeuValGlyLeuValTrpLeuValSerHisLeuSerG 120

Db 51325 ACTGGGGGTTCTGGGCTCTGCTGTGGTGTGCTTCACTGATCTGACCGCA 51266

QY 120 nArg 121

Db 51265 GCGG 51262

## RESULT 10

LOCUS

AR415427

DEFINITION

Sequence 3064 from patent US 6639063.

ACCESSION

AR415427

VERSION

AR415427.1 GI:40170537

KEYWORDS

Unknown.

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE

1 (bases 1 to 519)

AUTHORS

Edwards, J.-B.D.M., Jøbert, S. and Giordano, J.-Y.

TITLES

ESTs and encoded human proteins

JOURNAL

Patent: US 6639063-A 3064 28-OCT-2003;

FEATURES

Location/Qualifiers

source

1..519

/organism="unknown"

/mol\_type="genomic DNA"

## ORIGIN

Alignment Scores:

Pred. No.: 3,46e-33 Length: 519  
Score: 503.00 Matches: 99  
Percent Similarity: 99.00% Conservative: 0  
Best Local Similarity: 99.00% Mismatches: 1  
Query Match: 79.59% Indels: 1  
DB: 6 Gaps: 0

US-10-001-885-125 (1-121) x AR415427 (1-519)

QY 1 MetValArgIleuValAsnGlyGluIleValGlnAspAspProArgValArgThr 20

Db 219 ATGGTACGATCTGGCTATGAGGAGATCGTTCAGATGATGACCCCGAGTGGAGCT 278

QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40

Db 279 ACTACCCAGCCACCAAGAGTGCATCTCTGACAGAGCTTCTTCATAGGAGCTGCT 338

QY 41 AlaProProGlyIleProGlyProArgGlnGlnAlaGlyAlaGlyLeuGlyAlaI 60

Db 339 GCTCCCGCCAGG-GCTCTGCGCCCGCCCGACGACGAGGAGTGGAGTGGCT 397

QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80

Db 398 CAGTCCCGCTTCAATGACCTCAACCGGACGCTGTGTAAATGAGGCTTCCGAGTGGCAT 457

QY 81 LeuGlyAsnHisAlaValGluProValTrpSerIleLeuLeuPheLeuMetMet 100

Db 458 CTCGGCAACCATCTGTGAGCGGAGTGCATCTCTCTCTCTCTCTCATGATG 517

## RESULT 11

LOCUS

BD110980

DEFINITION

EST and encoded human protein.

ACCESSION

BD110980

CT 3205789

KEYWORDS JP 2002010789-A/3057.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. (bases 1 to 519)  
AUTHORS Edwards, J.B.D.M., Jobert, S., and Giordano, J.E.  
TITLE EST and encoded human protein  
JOURNAL Parent: JP 2002010789-A 3057 15-JAN-2002;  
GENSET CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002010789-A/3057  
PD 15-JAN-2002  
PF 07-AUG-2000 JP 2000280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N15/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N15/00, PC  
C12N15/00  
CC EST and encoded human protein  
FH Key Location/Qualifiers  
FT CDS Location/Qualifiers  
1..519  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,46e-33 Length: 519  
Score: 503.00 Matches: 99  
Percent Similarity: 99.00% Conservatave: 0  
Best Local Similarity: 99.00% Mismatches: 1  
Query Match: 79.59% Indels: 1  
DB: Gaps: 0  
US-10-001-885-125 (1-121) x BD110980 (1-519)  
QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspAppProArgValArgThr 20  
Db 219 ATGTGTGGGATCTTGGCCAAATGGGAAATCGTCAGATATGACGACCCCGAGAGAGACC 278  
QY 21 ThrThrglnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyIleGly 40  
Db 279 ACTACCCAGCCACAGAGAGTACCTTCGACAGAGCTTCTCATAGGGGCGCATGCT 338  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
Db 339 GCTCCCCAGAG-GGTCTGGGCCCCCGCAGCAGCAGCAGGCTGCMAGCTGGTGTGCT 397  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
Db 398 CAGTCCCCCTTCATGACCTCAACCGGACGCTGTGACATGGGCTTCCGAGTGGCAT 457  
QY 81 LeuGlyAsnHisAlaValAlaGluProValThrSerIleLeuLeuLeuPheLeuMetMet 100  
Db 458 CTCGGCAACCATGCTGTGAGAGCCGGTACCTCCATCTGCTCTCTCTGTCATCATG 517  
RESULT 12  
AL450311 165110 bp DNA linear PRI 12-JUL-2001  
LOCUS Human DNA sequence from clone Rpl1-343J3 on chromosome 10, complete  
DEFINITION sequence.  
ACCESSION AL450311  
VERSION AL450311.11 GI:14626972  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 165110)  
AUTHORS Howden, P.  
TITLE Direct Submission  
JOURNAL Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
COMMENT On Jul 8, 2001 this sequence version replaced gi:14575291.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em: EMBL; Sw:  
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; information on the WORMPEP  
database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 10, constructed by the Sanger Centre Chromosome 10  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr10  
Rpl1-343J3 is from the library RPl1-11.2 constructed by the group  
of Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pBACe3.6  
This sequence is the entire insert of clone Rpl1-343J3 The true  
left end of clone Rpl1-242G20 is at 139955 in this sequence. The  
true right end of clone Rpl1-404C6 is at 6588 in this sequence.  
FEATURES  
Source  
1..165110  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/clone="Rpl1-343J3"  
/clone\_11b="RPl1-11.2"  
7..147  
/note="TH1C repeat: matches 2..142 of consensus"  
151..293  
/note="MIR repeat: matches 2..148 of consensus"  
1136..1187  
/note="26 copies 2 mer gt 98% conserved"  
2960..3272  
/note="AluSg repeat: matches 1..310 of consensus"  
4144..4270  
/note="AluX repeat: matches 1..114 of consensus"  
4319..4494  
/note="AluX repeat: matches 118..293 of consensus"  
5662..6438  
/note="AluX repeat: matches 272..1095 of consensus"  
6916..8080  
/note="L1MEC repeat: matches 1168..2267 of consensus"  
9934..10087  
/note="MIR repeat: matches 25..184 of consensus"  
10353..10481  
/note="MIR repeat: matches 65..194 of consensus"  
11077..11372  
/note="AluX repeat: matches 1..294 of consensus"  
11600..11639  
/note="20 copies 2 mer tc 95% conserved"  
12479..12608  
/note="65 copies 2 mer at 90% conserved"  
13433..13876  
/note="MIR repeat: matches 1..466 of consensus"  
14372..14576

```
repeat_region /note="Alusg repeat: matches 1. .301 of consensus"  
14577. .14633  
/note="Alusg repeat: matches 243. .239 of consensus"  
repeat_region 14843. .14943  
/note="MER86 repeat: matches 9. .109 of consensus"  
repeat_region 15404. .15613  
/note="MIR repeat: matches 3. .213 of consensus"  
repeat_region 16626. .16705  
/note="MIR repeat: matches 60. .140 of consensus"  
repeat_region 16928. .16993  
/note="33 copies 2 mer gg 66% conserved"  
repeat_region 17476. .17569  
/note="MER81 repeat: matches 2. .114 of consensus"  
repeat_region 17719. .18069  
/note="LIMC4 repeat: matches 7617. .7977 of consensus"  
repeat_region 18117. .18345  
/note="MIR repeat: matches 7. .262 of consensus"  
repeat_region 19169. .19400  
/note="MIR repeat: matches 5. .239 of consensus"  
repeat_region 19436. .19627  
/note="LIMC5 repeat: matches 7728. .7917 of consensus"  
repeat_region 19642. .19716  
/note="MIR repeat: matches 48. .131 of consensus"  
repeat_region 19969. .20010  
/note="21 copies 2 mer tg 100% conserved"  
repeat_region 20623. .20704  
/note="41 copies 2 mer gt 85% conserved"  
repeat_region 20738. .20773  
/note="U2 repeat: matches 1. .36 of consensus"  
repeat_region 20802. .20863  
/note="MLT1J repeat: matches 1. .62 of consensus"  
repeat_region 20889. .21263  
/note="MLT1F repeat: matches 188. .541 of consensus"  
repeat_region 21463. .21618  
/note="MIR repeat: matches 46. .192 of consensus"  
repeat_region 22019. .22326  
/note="Alub repeat: matches 1. .306 of consensus"  
repeat_region 22381. .22564  
/note="U2 repeat: matches 2453. .2629 of consensus"  
repeat_region 22836. .23174  
/note="MLT1J repeat: matches 117. .413 of consensus"  
repeat_region 23215. .23346  
/note="MIR repeat: matches 48. .188 of consensus"  
repeat_region 23388. .23532  
/note="U2 repeat: matches 2097. .2230 of consensus"  
repeat_region 23533. .23637  
/note="Alub repeat: matches 1. .303 of consensus"  
repeat_region 23838. .24137  
/note="U2 repeat: matches 1754. .2097 of consensus"  
repeat_region 24291. .24581  
/note="Alusx repeat: matches 1. .300 of consensus"  
repeat_region 24653. .24850  
/note="MIR repeat: matches 1. .200 of consensus"  
misc_feature 26189. .26685  
/note="CpG island"  
/evidence=not_experimental  
29453. .30918  
/note="CpG island"  
/evidence=not_experimental  
31084. .31156  
/note="MIR repeat: matches 65. .138 of consensus"  
repeat_region 31618. .31876  
/note="Alub repeat: matches 29. .275 of consensus"  
repeat_region 32767. .32830  
/note="MIR repeat: matches 76. .139 of consensus"  
repeat_region 33050. .33178  
/note="43 copies 3 mer tcc 72% conserved"  
misc_feature 35112. .36201  
/note="CpG island"  
/evidence=not_experimental  
36551. .36604  
repeat_region
```

```
repeat_region /note="LIP15 repeat: matches 5480. .6157 of consensus"  
38265. .38561  
/note="Alusx repeat: matches 1. .300 of consensus"  
repeat_region 39214. .39423  
/note="MIR repeat: matches 3. .219 of consensus"  
repeat_region 39489. .39601  
/note="Alusx repeat: matches 1. .292 of consensus"  
repeat_region 40050. .40189  
/note="MIR repeat: matches 1. .144 of consensus"  
repeat_region 41057. .41290  
/note="MIR repeat: matches 8. .240 of consensus"  
repeat_region 41296. .41460  
/note="U2 repeat: matches 2569. .2730 of consensus"  
repeat_region 41506. .41944  
/note="U2 repeat: matches 1916. .2416 of consensus"  
repeat_region 42388. .42698  
/note="U2 repeat: matches 1448. .1779 of consensus"  
repeat_region 44193. .44579  
/note="TRIC repeat: matches 1. .371 of consensus"  
repeat_region 44600. .44956  
/note="LTR16A repeat: matches 90. .445 of consensus"  
repeat_region 45240. .45300  
/note="MER58A repeat: matches 37. .97 of consensus"  
repeat_region 45798. .45909  
/note="LTR41 repeat: matches 90. .192 of consensus"  
repeat_region 46826. .46871  
/note="23 copies 2 mer gt 97% conserved"  
repeat_region 50684. .50969  
/note="Alusg repeat: matches 1. .287 of consensus"  
repeat_region 50980. .51291  
/note="Alusx repeat: matches 1. .312 of consensus"  
repeat_region 52222. .52519  
/note="Alusx repeat: matches 1. .300 of consensus"  
repeat_region 54065. .54260  
/note="LIM4 repeat: matches 3865. .4055 of consensus"  
repeat_region 54261. .54432  
/note="PAM repeat: matches 2. .167 of consensus"  
repeat_region 54433. .54629  
/note="LIM4 repeat: matches 3652. .3865 of consensus"  
repeat_region 54648. .54862  
/note="LIR41 repeat: matches 11. .217 of consensus"  
repeat_region 54863. .55236  
/note="MLT1A1 repeat: matches 1. .365 of consensus"  
repeat_region 55237. .55700
```

Alignment Scores:

Pred. No.:	Length:	165110
Score:	472.00	Matches: 90
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	74.68%	Indels: 0
DB:	9	Gaps: 0

US-10-001-885-125 (1-121) x AL450311 (1-165110)

```
QY 32 GlnSerPheAsnArgIyHnIslgYAlAProProglYgIyProglYProArgIngn 51  
Db 9069 CAGAGCTTCTTCAACAGGGGCGCATGTGCTCCCGAGGGGGTCTGGGCCCGCAGAG 90128  
QY 52 GlnAlaGlyAlAArgLeuGlyAlaAlaGlnSerProPheAsnArgIngn 71  
Db 90129 CAGGACAGGCGGAGGCTGCTGCTGCTCCCTTCAAGACTTCAACCGGAGCTG 90188  
QY 72 ValAsnMetGlyPheProGlnTrpHisLeuGlyAsnHisAlaValGlnProValThSer 91  
Db 90189 GTCACATAGGCTTTCGACAGTGCATCTTGGCAACACATGCTGAGCGGCTGACTCC 90248  
QY 92 IleLeuLeuPheLeuLeuMetLeuLeuGlyValArgGlyLeuLeuLeuValGlyLeu 111  
Db 90249 ATCTGCTCTCTCTCTGCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 90308
```

Db 90309 GTCTACTGTGTCACCTGATGACGG 90338

RESULT 13  
AC021954/c  
LOCUS  
DEFINITION  
AC021954  
AC021954  
AC021954.3 GI:7417809  
HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
1 (bases 1 to 173341)  
Britten, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 10, clone RP11-57E12  
2 (bases 1 to 173341)  
Britten, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckert, R., Beda, F., Boguslavsky, L., Bouckhalter, B., Brown, A., Burkett, G., Castle, A., Choapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, K., DeRellano, K., Dewar, K., Domino, M., Doyle, M., Fennell, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardina, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howard, T., Johnson, R., Jones, C., Kamm, L., Karatas, A., Klein, J., Lander, T., Lehoczy, J., Levine, R., Liu, G., Locke, K., MacDonald, P., Margulis, N., McEwan, P., McGurk, A., McKernan, K., McHeesters, R., Meldrum, J., Menus, L., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K., Piers, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theadore, J., Titrill, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Submitted (22-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 173341)  
Britten, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Bouckhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choapel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, K., DeRellano, K., Dewar, K., Diaz, J. S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardina, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howard, T., Johnson, R., Jones, C., Kamm, L., Karatas, A., Klein, J., Lehoczy, K., Lamazares, R., Lander, T., Lehoczy, J., Levine, R., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McHeesters, R., Meldrum, J., Menus, L., Morrow, J., Miranda, C., Mings, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testa, S., Theadore, J., Titrill, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 5, 2000 this sequence version replaced gi:6984451.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997) RepeatMasker.html  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
Center: Whitehead Institute/MIT Center for Genome Research  
Center code: WIRB  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

Project Information  
Center project name: L5931  
Center clone name: 57 E.12  
Summary Statistics  
Sequencing vector: M13, M77815, 100% of reads  
Chemistry: Dye-terminator Big, Dye, 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 161190 bases at least Q40  
Consensus quality: 166837 bases at least Q30  
Consensus quality: 168995 bases at least Q20  
Insert size: 176000; agarose-fp  
Insert size: 171041; sum-of-contigs  
Quality coverage: 3.7 in Q20 bases; sum-of-contigs  
Quality coverage: 3.8 in Q20 bases; sum-of-contigs  
NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  
1 1574: contig of 1574 bp in length  
1575 1674: gap of 100 bp  
1675 3043: contig of 1389 bp in length  
3044 3143: gap of 100 bp  
3144 5889: contig of 2746 bp in length  
5890 5989: gap of 100 bp  
5990 8979: contig of 2990 bp in length  
8980 9080: gap of 100 bp  
9080 13674: contig of 4595 bp in length  
13675 13774: gap of 100 bp  
13775 18831: contig of 5057 bp in length  
18832 18931: gap of 100 bp  
18932 23526: contig of 4595 bp in length  
23527 23627: gap of 100 bp  
23627 27386: contig of 3760 bp in length  
27387 27486: gap of 100 bp  
27487 32572: contig of 5086 bp in length  
32573 32673: gap of 100 bp  
32673 38632: contig of 5960 bp in length  
38633 38732: gap of 100 bp  
38733 43735: contig of 5003 bp in length  
43736 43835: gap of 100 bp  
43836 49020: contig of 5185 bp in length  
49021 49120: gap of 100 bp  
49121 53660: contig of 4540 bp in length  
53661 53760: gap of 100 bp  
53761 59544: contig of 5784 bp in length  
59545 59644: gap of 100 bp  
59645 66395: contig of 6751 bp in length  
66396 66496: gap of 100 bp  
66497 74596: contig of 8201 bp in length  
74597 74796: gap of 100 bp  
74797 81326: contig of 6530 bp in length  
81327 81426: gap of 100 bp  
81427 89446: contig of 8020 bp in length  
89447 89546: gap of 100 bp  
89547 96459: contig of 6913 bp in length  
96460 96559: gap of 100 bp  
96560 106498: contig of 9939 bp in length  
106499 106598: gap of 100 bp  
106599 116986: contig of 10388 bp in length  
116987 117086: gap of 100 bp  
117087 128890: contig of 11804 bp in length  
128891 128990: gap of 100 bp  
128991 147290: contig of 18300 bp in length  
147291 147390: gap of 100 bp  
147391 173341: contig of 25951 bp in length.  
Location/Qualifiers  
1..173341  
/organism="Homo sapiens"

```

misc_feature      /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /chromosome="10"
                  /map="10"
                  /clone_id="RP11-57E12"
                  /clone_lib="RP11-11 Human Male BAC"
                  1..1574
misc_feature      /note="assembly_fragment"
                  1675..3043
misc_feature      /note="assembly_fragment"
                  3144..5889
misc_feature      /note="assembly_fragment"
                  5990..8979
misc_feature      /note="assembly_fragment"
                  9080..13674
misc_feature      /note="assembly_fragment"
                  13775..18831
misc_feature      /note="assembly_fragment"
                  18932..23526
misc_feature      /note="assembly_fragment"
                  23627..27386
misc_feature      /note="assembly_fragment"
                  clone_end:SP6
                  vector_side:left"
                  27487..32572
misc_feature      /note="assembly_fragment"
                  32673..38632
misc_feature      /note="assembly_fragment"
                  38733..43735
misc_feature      /note="assembly_fragment"
                  43836..49020
misc_feature      /note="assembly_fragment"
                  49121..53660
misc_feature      /note="assembly_fragment"
                  53761..59544
misc_feature      /note="assembly_fragment"
                  59645..66395
misc_feature      /note="assembly_fragment"
                  66496..74696
misc_feature      /note="assembly_fragment"
                  74797..81326
misc_feature      /note="assembly_fragment"
                  81427..89446
misc_feature      /note="assembly_fragment"
                  89547..96459
misc_feature      /note="assembly_fragment"
                  96560..106498
misc_feature      /note="assembly_fragment"
                  106599..116986
misc_feature      /note="assembly_fragment"
                  117087..128890
misc_feature      /note="assembly_fragment"
                  clone_end:P7
                  vector_side:right"
                  128991..147290
misc_feature      /note="assembly_fragment"
                  147391..173341
misc_feature      /note="assembly_fragment"

Alignment Scores:
Pred. No.:      5,47e-28      Length:      173341
Score:          472.00      Matches:      90
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    74.68%      Indels:      0
DB:             2      Gaps:      0

US-10-001-885-125 (1-121) x AC021954 (1-173341)
QY      32 GlnSerPheAsnArgGlyHisGlyAlaProProGlyGlyProGlyProArgGlnGln 51
DB      CAGAGCTTCTTCAATAGAGGGCCATGTCCTCTGCGCCCGCCACAG 161030
```

```

QY      52 GlnAlaGlyAlaAglGlyAlaAlaGlnSerProPheAsnAspLeuAsnArgGlnLeu 71
DB      CAGGACAGGTGCGAGGCTGGGCTGCTCAGTCCCTTCATGACCTCAACCGGACGCTG 160970
QY      72 ValAsnMetGlyPheProGlnTrpHisLeuGlyAsnHisAlaValGluProValThrSer 91
DB      GTGAACATGCGGCTTTCCCGCAGATGGCATCTCGGCAACATGCTGTGGAGCCGCTACCTCC 160910
QY      92 IleLeuLeuLeuPheLeuLeuMetMetLeuGlyValAlaGlyLeuLeuValGlyLeu 111
DB      ATCCGTGCTCCTCTCTCTCTCTCATGATGCTGTGCTGCTGCGCTCTCTCTGTTGGCTT 160850
QY      112 ValTyrLeuValSerHisLeuSerGlnArg 121
DB      GTCTACCTGGTGTCCCACTGATGATGCGG 160820
Db      160849 GTCTACCTGGTGTCCCACTGATGATGCGG 160820

RESULT 14
AC127417      215050 bp      DNA      linear      HTG 19-OCT-2002
LOCUS      Mus musculus chromosome UNK clone RP23-459M2, WORKING DRAFT
DEFINITION      SEQUENCE, 7 unordered pieces.
AC127417
AC127417.2 GI:24137619
VERSION      AC127417.2
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE      The sequence of Mus musculus clone
JOURNAL      Unpublished
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE      2 (bases 1 to 215050)
JOURNAL      Direct Submission
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE      Submitted (19-OCT-2002) Genome Sequencing Center, 4444 Forest Park
JOURNAL      Parkway, St. Louis, MO 63108, USA
COMMENT      On Oct 19, 2002 this sequence version replaced gi:21759524.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
Project information
Center project name: M_BA0459M02
----- Summary Statistics -----
Sequencing vector: M13, 0%
Sequencing vector: plasmid: 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly Program: Phrap; Version 0.990319
Consensus quality: 212279 bases at least Q40
Consensus quality: 213137 bases at least Q30
Consensus quality: 213590 bases at least Q20
Insert size: 19200; agarose-fp
Insert size: 217944; sum-of-contigs
Quality coverage: 11.89 in Q20 bases; agarose-fp
Quality coverage: 10.54 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
```

```

* be preserved.
* 1 4227: contig of 4227 bp in length
* 4228 4327: gap of unknown length
* 4328 16589: contig of 12262 bp in length
* 16590 16689: gap of unknown length
* 16690 27318: contig of 10629 bp in length
* 27319 27418: gap of unknown length
* 27419 48797: contig of 22379 bp in length
* 48798 48897: gap of unknown length
* 48898 79378: contig of 29481 bp in length
* 79379 132595: gap of unknown length
* 132596 132695: contig of 5117 bp in length
* 132696 215050: gap of unknown length
* 215051 82355: contig of 82355 bp in length.

FEATURES
    source
        1..215050
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10030"
            /chromosome="XNK"
            /clone="RP23-459K2"
        1..4227
            /note="assembly_name:Contig31"
        misc_feature
            4328..16589
                /note="assembly_name:Contig32"
        misc_feature
            16690..27318
                /note="assembly_name:Contig33"
        misc_feature
            27419..48797
                /note="assembly_name:Contig34"
        misc_feature
            48898..79378
                /note="assembly_name:Contig35"
        misc_feature
            79379..132595
                /note="assembly_name:Contig36"
        misc_feature
            132596..215050
                /note="assembly_name:Contig37"

ORIGIN
Alignment Scores:
Pred. No.:      8,16e-27      Length:      215050
Score:          459.00      Matches:      87
Percent Similarity: 97.78%      Conservative: 1
Best Local Similarity: 96.67%      Mismatches: 2
Query Match:    72.63%      Indels:      0
Gaps:           0

US-10-001-885-125 (1-121) X AC127417 (1-215050)

Qy      32 GlnserPhepheaanaaglyghisgylalarpProctlygylProglyProargIngn 51
Db      181230 CAGGCGCTTTTTCACAGAGCGCACGCGCACCTTCAGGCGGCCCTCGAACCCCGCACGAG 181171

Qy      52 GlnalaglyalaaglygylalalaglnserProphaeanaapleuanAarglnleu 71
Db      181170 CAGGCAAGGTGCCGACACTGGGCTGCCCAANTCTCTTCAGTACCTGGAACCGGCAAGCTG 181111

Qy      72 ValamMetglyPhepProglntTPhleuglyanhsisalavalgluprovalThrsr 91
Db      181110 GTGAACATGGGCTTCCACACATGCGACCTTGGGAAACACGCTGGGAACTGTGACCTCC 181051

Qy      92 IleuenuleuPheleuLeuMetMetleuglyValaargglyleuLeuValglyleu 111
Db      181050 ATCTCTCTGCTCTCTCTGCTTATATAGCTCGGCGCTTCGAGGCTCTCTGCTTGTGGGCTTG 180991

Qy      112 ValtyrleuValserHisleuSerGlnary 121
Db      180990 GTCTACCTGGTGTCTCACTGAGTCAGCGGG 180961

RESULT 15
AC127817/c AC127817 258815 bp DNA linear HTG 19-NOV-2002
LOCUS      Rattus norvegicus clone CH230-259G16, WORKING DRAFT SEQUENCE, 3
DEFINITION unorderd pieces.
ACCESSION AC127817

```

```

VERSION
AC127817.3 GI:25077905
KEYWORDS
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 258815)
REFERENCE
Muzny,D,Marie,, Metzker,M, Lee,, Abramson,S,, Adams,C,, Alder,J,,
Allen,C,, Allen,H,, Alperbrooks,S,, Amin,A,, Angiulano,D,,
Anyalabeche,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,,
Baldridge,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F,,
Biswal,K,, Blair,J,, Blankenburg,K,, Blyth,P,, Brown,M,,
Bryant,N,, Buhay,C,, Burch,P,, Butrell,K,, Calderon,E,,
Cardenas,V,, Carter,K,, Cavazos,I,, Ceasar,H,, Center,A,,
Chacko,J,, Chavez,D,, Chen,G,, Chen,R,, Chen,Y,, Chen,Z,, Chu,J,,
Cleveland,C,, Cockrell,R,, Cox,C,, Coyle,M,, Cree,A,, D'Gouza,L,,
Davila,M,L,, Davis,C,, Davy-Carroll,L,, De Andrea,C,, Dedrich,D,,
Delgado,O,, Denson,S,, Deramo,C,, Ding,Y,, Din,H,, Divya,K,,
Draper,H,, Dugan-Rocha,S,, Dunn,A,, Durbin,K,, Duval,B,, Eaves,K,,
Egan,A,, Escotto,M,, Eugene,C,, Evans,C,A,, Falls,T,, Fan,G,,
Fernandez,S,, Finley,M,, Flagg,N,, Forbes,L,, Foster,M,, Foster,P,,
Frazer,C,M,, Gabisi,A,, Ganta,R,, Garcia,A,, Garner,T,, Garza,M,,
Gebregorgis,E,, Geer,K,, Gill,R,, Grady,M,, Gueria,W,, Guevara,W,,
Guraratne,P,, Haaland,W,, Hamil,C,, Hamilton,C,, Hamilton,K,,
Harvey,Y,, Havlak,P,, Hawes,A,, Henderson,N,, Hernandez,J,,
Hernandez,R,, Hines,S,, Hladun,S,L,, Hodgson,A,, Hognes,M,,
Hollins,B,, Howells,S,, Hulik,S,, Hune,V,, Idlebird,D,, Jackson,A,,
Jackson,L,, Jacob,L,, Jiang,H,, Johnson,B,, Johnson,R,, Jolivet,A,,
Karpaty,S,, Kelly,S,, Kelly,S,, Khan,Z,, King,L,, Kovar,C,,
Kowis,C,, Kraft,C,L,, Lebow,H,, Levan,J,, Lewis,L,, Li,Z,, Liu,J,,
Liu,J,, Liu,M,, Liu,Y,, London,P,, Longacre,S,, Lopez,J,,
Lorenshewa,L,, Loulseged,H,, Lozano,R,J,, Lu,X,, Ma,J,,
Maheshwari,M,, Mahindaratne,M,, Mahmoud,M,, Malloy,K,, Mangum,A,,
Mangum,B,, Mapa,P,, Martin,K,, Martin,R,, Martinez,E,,
Mawhney,S,, McLeod,M,P,, McNeill,T,Z,, Meenen,E,,
Milosavljevic,A,, Miner,G,, Minja,E,, Montemayor,J,, Moore,S,,
Morgan,M,, Morris,K,, Morris,S,, Mundtaser,M,, Murphy,M,, Nair,L,,
Nankervels,C,, Neal,D,, Newton,N,, Ngunyana,N,, Norris,S,,
Nwackeleme,O,, Okonou,G,, Olarnunatgon,A,, Pal,S,, Parks,K,,
Pasternak,S,, Paul,H,, Perez,A,, Perez,L,, Pfannkoch,C,,
Pfeiffer,F,, Poindexter,A,, Popovic,D,, Primus,E,, Pu,L,, Li,,
Puzo,M,, Quiroz,J,, Rachlin,E,, Reeves,K,, Regier,M,A,, Reigh,R,,
Reilly,B,, Reilly,M,, Ren,Y,, Reuter,M,, Richards,S,, Riggs,F,,
Rivers,C,, Rodkey,T,, Rojas,A,, Rose,M,, Rose,R,, Ruiz,S,D,,
Sanders,M,, Savery,G,, Scherer,S,, Scott,G,, Shatman,S,, Shen,H,,
Sherry,J,, Shvartsbeyn,A,, Sison,I,, Sitter,C,D,, Smas,D,,
Sneed,A,, Sodergren,E,, Song,X,-Z,, Sorelle,R,, Sosa,J,,
Steinle,M,, Strong,R,, Sutton,A,, Svatek,A,, Taber,P,, Taylor,C,,
Taylor,T,, Thomas,N,, Thomas,S,, Tingey,A,, Trejos,P,, Usman,K,,
Valas,R,, Vera,V,, Villasana,D,, Waldron,L,, Walker,B,, Wang,J,,
Wang,Q,, Wang,S,, Warren,J,, Warren,R,, Wei,X,, White,F,,
Williams,G,, Willison,R,, Wleczek,R,, Wooden,H,, Worley,K,,
Wright,D,, Wright,R,, Wu,J,, Yakub,S,, Yen,J,, Yoon,L,, Yoon,V,,
Yu,F,, Zhang,Y,, Zhou,J,, Zhou,X,, Zhao,S,, Dunn,D,, von
Niederhusern,A,, Weiss,R,, Smith,D,R,, Holt,R,A,, Smith,H,O,,
Weinstock,G,, and Gibbs,R,A.
Direct Submission
Unpublished
2 (bases 1 to 258815)
REFERENCE
Worley,K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 258815)
REFERENCE
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 19, 2002 this sequence version replaced gi:23912578.
The sequence in this assembly is a combination of BAC based reads

```

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GZXS  
Center clone name: CH230-259G16  
----- Summary Statistics

Assembly program: Phrap; version 0.990329  
Consensus quality: 224747 bases at least Q40  
Consensus quality: 227981 bases at least Q30  
Consensus quality: 229752 bases at least Q20  
Estimated insert size: 228243; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  
\* 1 255979: contig of 255979 bp in length  
\* 255980 256079: gap of unknown length  
\* 256080 257349: contig of 1270 bp in length  
\* 257350 257449: gap of unknown length  
\* 257450 258815: contig of 1366 bp in length.

FEATURES  
source  
1..258815  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-259G16"

misc\_feature  
1..1609  
/note="wgs end extension  
clone\_end:5p6"

misc\_feature  
/note="clone boundary  
complement(2865..3659)  
clone\_end:5p6  
site:  
end\_sequence:BZ169379"

misc\_feature  
224098..225157  
/note="clone boundary  
clone\_end:17  
site:  
end\_sequence:BZ169378"

# ORIGIN

Alignment Scores:  
Pred. No.: 8,056-26 Length: 258815  
Score: 448.00 Matches: 85  
Percent Similarity: 95.56% Conservative: 1  
Best Local Similarity: 94.44% Mismatches: 4  
Query Match: 70.89% Indels: 0  
DB: 2 Gaps: 0

US-10-001-885-125 (1-121) x AC127817 (1-258815)

QY 32 GlnserPhePheAsnArgGlyHisGlyValProProGlyGlyProGlyProArgGlnGln 51  
Db 220380 CAGGGCTTTTCAACAGAGGCCACAGTCACCTCCAGGGGGCCCTGGACCCACAGCAG  
QY 52 GlnAlaGlyAlaArgLeuGlyValAlaGlnserProPheAsnAspLeuAsnArgGlnLeu 71  
Db 220320 CAGGACAGGCGCCGACTGGGTGCTCCCAATCTCTTCAAGTACCTGAACCGGAGCTG  
QY 72 ValAsnMetGlyPheProGlnTrpHisLeuGlyAsnHisAlaValGluProValThrSer 91  
Db 220260 GTGAACATGGGCTTCCACAAATGGCATTTGGGAACACGCGTAGACCTGTGACCTCC  
QY 92 IleLeuLeuLeuPheLeuLeuValMetLeuGlyValAlaArgGlyLeuLeuValGlyLeu 111  
Db 220200 ATCCTCTCTCTCTCTCTCTCTCTCATGATGCTAGGGGTTCTGTCCTCTCTCTGAGGCTG 220141  
QY 112 ValTyrLeuValSerHisLeuSerGlnArg 121  
Db 220140 GTCTACTGCTGTCTCTCACTGAGCCAGCGG 220111

Search completed: August 22, 2004, 09:14:41  
Job time : 2208 secs



OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2004, 07:09:25 ; Search time 339 Seconds

(without alignments)  
1516.318 Million cell updates/sec

Title: US-10-001-885-125

Perfect score: 632

Sequence: 1 MVRILANGELVQDDPRVTRT.....GVRGLLVGLVIVSHLSQR 121

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 segs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp  
-O=/gen2.1/USPTO.spool.p/US10001885/runat.17082004.151717.25892/app.query.fasta\_1.263  
-DB=Geneseq 291n04 -OPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOFEFT=0 -UNITS=bits -START=1 -END=1 -MATRIX=dl0sune2 -TRANS=human40.cdt  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10001885.OCGN.1.1.708/runat.17082004.151717.25892 -NCPU=6 -ICPU=3  
-NO MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAUS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29jan04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002s:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	632	100.0	365	7	ABX92251 Human ova
2	632	100.0	562	6	ABZ11212 Human pol
3	632	100.0	1049	2	AAV59558 Human sec
4	632	100.0	1049	6	AB573545 Human cdn
5	632	100.0	1049	8	ACD82688 CDNA sequ
6	632	100.0	1108	6	ABL90057 Human pol
7	632	100.0	1192	6	ABN87820 Human ova
8	577	91.3	550	4	AAF82505 Human bre

C	9	434	68.7	1165	5	AA64820 DNA encod
C	10	113	17.9	517	7	ABX74673 Human CDN
C	11	110.5	17.5	1110	9	ABE07153 Novel cod
C	12	110.5	17.5	2999	7	AB557565 Human SEC
C	13	107	16.9	2668	7	ACA03995 Human SEC
C	14	89	14.1	4177	5	AA545077 CDNA down
C	15	89	14.0	4839	5	ABV30270 Human enco
C	16	88.5	14.0	2172	4	AB51456 Human pro
C	17	88.5	14.0	2172	7	ACA19536 Human pro
C	18	88.5	14.0	2172	8	ACC59400 Microbial
C	19	88	13.9	1122	7	ADA71003 Rice gene
C	20	86.5	13.7	494	8	ACH43729 Human foe
C	21	86.5	13.7	1251	3	AAA62022 Hydrophob
C	22	86.5	13.7	1617	6	AAA62032 Hydrophob
C	23	86.5	13.7	2417	6	ABX34832 Human CDN
C	24	86.5	13.7	2487	4	AAH14983 Human CDN
C	25	86.5	13.7	2562	4	AAH14983 Human CDN
C	26	85	13.4	8031	4	AAH07167 Human mem
C	27	85	13.4	8066	4	AAH07167 Human rep
C	28	84.5	13.4	1024	7	ABZ83399 Toxicolog
C	29	84.5	13.4	1170	7	ACD05761 CDNA enco
C	30	84.5	13.4	1359	4	AA94501 Human hya
C	31	84.5	13.4	1484	6	ABT11180 Human sec
C	32	84.5	13.4	1548	5	AA582715 DNA encod
C	33	84.5	13.4	1550	2	AA791346 Human HO-
C	34	84.5	13.4	1550	2	AA83838 Human hae
C	35	84.5	13.4	1550	3	AA65120 Human hae
C	36	84.5	13.4	1550	3	AA293230 Human Hem
C	37	84.5	13.4	1550	6	ABE2836 Human hae
C	38	84.5	13.4	1550	6	ABE2836 Breast ca
C	39	84.5	13.4	1550	6	ABE63105 Human NF-
C	40	84.5	13.4	1550	6	ABQ91974 Gene #368
C	41	84.5	13.4	1558	6	ABN97191 Human hya
C	42	84.5	13.4	1558	4	AA94511 Human tun
C	43	84.5	13.4	1907	7	ADD29600 Human cod
C	44	84.5	13.4	2259	9	ADA53012 Human cod
C	45	84.5	13.4	2772	9	ADB63122 Human CDN

ALIGNMENTS

RESULT 1  
ABX92251 standard; CDNA; 365 BP.

AC	XX	ABX92251;	
AC	XX	ABX92251;	
DT	08-MAY-2003	(first entry)	
DE	Human ovarian specific nucleic acid DEX0310_65.		
XX	Human; gene; ss; ovarian specific nucleic acid; OSNA; ovarian cancer;		
KW	non-cancerous ovarian disease; gene therapy; vaccine; cytostatic;		
KW	gynaecological.		
OS	Homo sapiens.		
XX	WO200292785-A2.		
PN	21-NOV-2002.		
PD	13-FEB-2002; 2002WC-US022271.		
PF	13-FEB-2001; 2001US-0268290P.		
PR	15-FEB-2001; 2001US-026834P.		
PA	(DIAD-) DIADEXUS INC.		
PI	Salceda S, Macina RA, Hu P, Recipon H, Karra K, Caferkey R;		
PI	Sun Y, Liu C,		
XX	WPI; 2003-120677/11.		
DR	P-Psdb; ABU61063.		

XX New isolated OSNA nucleic acid and encoded polypeptide, useful for  
PT identifying, diagnosing, monitoring, staging, imaging and treating  
PT ovarian cancer and non-cancerous diseases in ovarian tissues.

PS Claim 1; Page 185-186; 224pp; English.

XX The invention relates to a new isolated nucleic acid, termed ovarian  
CC specific nucleic acid (OSNA) comprising: (a) a nucleic acid molecule that  
CC encodes any of 53 fully defined protein sequences appearing as AB061018-  
CC AB061070 (termed ovarian specific proteins, OSP); (b) any of 76 fully  
CC defined nucleotide sequences appearing as ABX92187-ABX92262; or (c) a  
CC sequence having at least 60% sequence identity to the nucleic acid  
CC molecule of (a) or (b). Also included are a method for determining the  
CC presence of an ovary specific nucleic acid (OSNA) in a sample, a vector  
CC comprising an OSNA, a host cell comprising the vector, an isolated OSP  
CC polypeptide, an anti-OSP antibody or fragment, a method for determining  
CC the presence of an ovary specific protein in a sample and a vaccine  
CC comprising an OSP or OSNA. The methods and compositions of the present  
CC invention are useful for identifying, diagnosing, monitoring, staging,  
CC imaging and treating ovarian cancer and non-cancerous disease in ovary  
CC tissue. The present sequence is an OSNA of the invention

XX Sequence 365 BP; 60 A; 117 C; 108 G; 80 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.:	2,24e-51	Length:	365
Score:	632.00	Matches:	121
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-10-001-885-125 (1-121) x ABX92251 (1-365)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 1 ATGGTGGCGATCTTGGCCAAATGGGGAATCGGACGATGACGCCCGAGTAAAGACC 60  
QY 21 ThrThgInpProArgGlySerIleProArgInserPheAsnArgGlyHisGly 40  
DB 61 ACTACCCAGCCCAAGAGGTRGCACTTCTCGACAGAGCTTCTCAATAGGGCCCATGCT 120  
QY 41 AlaProPheGlyIleProGlyProArgGlnGlnGlnAlaGlyValArgLeuGlyValAla 60  
DB 121 GGTCCCCCGGGGGTCTGTGCCCCCGCAGCAGCAGGAGTGCAGGCTGGGTGCTGCT 180  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 181 CAGTCCCTTAAAGACCTCAACCGGCACTGTGTAACATGGGCTTCCCGCATGGGCAT 240  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuLeuPheLeuLeuMet 100  
DB 241 CTCGGCAACCATGCTGTGGAGCCGGTGAACCTCCATCTGCTCTCTCATGATG 300  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTrpLeuValSerHisLeuSerGln 120  
DB 301 CTTGGTGTTCGGGCTCTCTGTTGTGCTTGTCTACCTGGTGTCCACCATGATGAC 360  
QY 121 Arg 121  
DB 361 CCG 363

RESULT 2  
AB211212  
ID AB211212 standard; cDNA; 562 BP.

AC AB211212;

XX 20-JAN-2003 (first entry)

KW Human; genome mapping; gene therapy; food supplement; virus; fungus;  
KW cell-proliferative disorder; neurodegenerative disease; bacterial;  
KW Parkinson's disease; Alzheimer's disease; autoimmune disease; bacterial;  
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;  
KW arthritis; cytostatic; immunomodulator; nocitropic; neuroprotective;  
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;  
KW haemostatic; valnerary; fungicide; antibacterial; virucide; protozoacide;  
KW antiarthritic; gene; ss.

XX Homo sapiens.

XX WO200270539-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US005095.

XX 05-MAR-2001; 2001US-00799451.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QH, Ren F,  
PI Xue HJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,  
PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2002-759812/82.

XX P-PSDB; ABP68995.

PT New polynucleotides comprising sequences assembled from expressed

PT sequence tags (ESTs), useful for treating cell-proliferative,  
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet  
PT or coagulation disorders.

XX Claim 1; SEQ ID NO 94; 1012pp + Sequence Listing; English.

CC The invention relates to an isolated polynucleotide (I) comprising a  
CC nucleotide sequence selected from any of 948 sequences (AB21119-  
CC AB212066) or their mature protein coding portion, active domain coding  
CC protein or complementary sequences. The polynucleotides are useful for  
CC identifying expressed genes or for physical mapping of human genome. The  
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
CC markers, as a food supplement, for generating antibodies, in medical  
CC imaging, screening and diagnostic assays and for treating cell-  
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, liver  
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was published in electronic format  
CC directly from WIFO at [http://wipo.int/pub/published\\_pat\\_sequences](http://wipo.int/pub/published_pat_sequences)

XX Sequence 562 BP; 95 A; 176 C; 174 G; 117 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.:	3.73e-51	Length:	562
Score:	632.00	Matches:	121
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-001-885-125 (1-121) x AB211212 (1-562)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 146 ATGGTGGCGATCTTGGCCAAATGGGGAATCGGACGATGACGCCCGAGTAAAGACC 205

QY 21 ThrThgInpProArgGlySerIleProArgInserPheAsnArgGlyHisGly 40

DB 206 ACTACCCAGCCCAAGAGGTRGCACTTCTCGACAGAGCTTCTCAATAGGGCCCATGCT 265

Db 266 GCTCCCCAGGCGGCTCTCCGCCCCCGCCAGCAGCAGCAGGCTGCGTGCCT 325  
QY 61 GlnSerProPheAnaapLeuAnaArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
Db 326 CAGTCCCCCTTCATGACCTTACACCGGACGCTGTGACATGGACCTTCGCGATGGCAT 385  
QY 81 LeuGlyAsnHisAlaValAlaGluProValThrSerIleLeuLeuPheLeuMetMet 100  
Db 386 CTCGCAACCATGCTGTGGAGCCGCGTACCTCCATCTGCTCTCTCCGTCATCATG 445  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValIleSerHisLeuSerGln 120  
Db 446 CTTCGTGTCTGGCCTCTCTCTCTGCTGTGGCCTTGTCTTACCTGTGTCACCTGAGTCAG 505  
QY 121 Arg 121  
Db 506 CGG 508  
RESULT 3  
AAVS5558  
ID AAVS5558 standard; DNA, 1049 BP.  
AAVS5558;  
AC AAVS5558;  
XX  
XX  
DT 06-JAN-1999 (first entry)  
XX  
DE Human secreted protein gene 48 clone HECAT174.  
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX  
OS Homo sapiens.  
XX  
XX W09839448-A2.  
PD 11-SEP-1998.  
XX  
XX  
PF 06-MAR-1998; 98MO-US004493.  
XX  
PR 07-MAR-1997; 97US-0038621P.  
PR 07-MAR-1997; 97US-0040161P.  
PR 07-MAR-1997; 97US-0040162P.  
PR 07-MAR-1997; 97US-0040163P.  
PR 07-MAR-1997; 97US-0040333P.  
PR 07-MAR-1997; 97US-0040334P.  
PR 07-MAR-1997; 97US-0040336P.  
PR 07-MAR-1997; 97US-0040626P.  
PR 11-APR-1997; 97US-0043311P.  
PR 11-APR-1997; 97US-0043312P.  
PR 11-APR-1997; 97US-0043313P.  
PR 11-APR-1997; 97US-0043314P.  
PR 11-APR-1997; 97US-0043315P.  
PR 11-APR-1997; 97US-0043568P.  
PR 11-APR-1997; 97US-0043569P.  
PR 11-APR-1997; 97US-0043576P.  
PR 11-APR-1997; 97US-0043578P.  
PR 11-APR-1997; 97US-0043580P.  
PR 11-APR-1997; 97US-0043669P.  
PR 11-APR-1997; 97US-0043670P.  
PR 11-APR-1997; 97US-0043671P.  
PR 11-APR-1997; 97US-0043672P.  
PR 11-APR-1997; 97US-0043674P.  
PR 23-MAY-1997; 97US-0047492P.  
PR 23-MAY-1997; 97US-0047500P.  
PR 23-MAY-1997; 97US-0047501P.  
PR 23-MAY-1997; 97US-0047502P.

PR 23-MAY-1997; 97US-0047503P.  
PR 23-MAY-1997; 97US-0047581P.  
PR 23-MAY-1997; 97US-0047582P.  
PR 23-MAY-1997; 97US-0047583P.  
PR 23-MAY-1997; 97US-0047584P.  
PR 23-MAY-1997; 97US-0047585P.  
PR 23-MAY-1997; 97US-0047586P.  
PR 23-MAY-1997; 97US-0047587P.  
PR 23-MAY-1997; 97US-0047588P.  
PR 23-MAY-1997; 97US-0047589P.  
PR 23-MAY-1997; 97US-0047590P.  
PR 23-MAY-1997; 97US-0047592P.  
PR 23-MAY-1997; 97US-0047593P.  
PR 23-MAY-1997; 97US-0047594P.  
PR 23-MAY-1997; 97US-0047595P.  
PR 23-MAY-1997; 97US-0047596P.  
PR 23-MAY-1997; 97US-0047597P.  
PR 23-MAY-1997; 97US-0047598P.  
PR 23-MAY-1997; 97US-0047599P.  
PR 23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047616P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047619P.  
PR 23-MAY-1997; 97US-0047632P.  
PR 23-MAY-1997; 97US-0047633P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 13-JUN-1997; 97US-0048974P.  
PR 08-JUL-1997; 97US-0051926P.  
PR 16-JUL-1997; 97US-0051927P.  
PR 18-AUG-1997; 97US-0055724P.  
PR 22-AUG-1997; 97US-005630P.  
PR 22-AUG-1997; 97US-005631P.  
PR 22-AUG-1997; 97US-005632P.  
PR 22-AUG-1997; 97US-005635P.  
PR 22-AUG-1997; 97US-005637P.  
PR 22-AUG-1997; 97US-005662P.  
PR 22-AUG-1997; 97US-005664P.  
PR 22-AUG-1997; 97US-0056845P.  
PR 22-AUG-1997; 97US-0056862P.  
PR 22-AUG-1997; 97US-0056864P.  
PR 22-AUG-1997; 97US-0056872P.  
PR 22-AUG-1997; 97US-0056874P.  
PR 22-AUG-1997; 97US-0056875P.  
PR 22-AUG-1997; 97US-0056876P.  
PR 22-AUG-1997; 97US-0056877P.  
PR 22-AUG-1997; 97US-0056878P.  
PR 22-AUG-1997; 97US-0056879P.  
PR 22-AUG-1997; 97US-0056880P.  
PR 22-AUG-1997; 97US-0056881P.  
PR 22-AUG-1997; 97US-0056882P.  
PR 22-AUG-1997; 97US-0056884P.  
PR 22-AUG-1997; 97US-0056885P.  
PR 22-AUG-1997; 97US-0056887P.  
PR 22-AUG-1997; 97US-0056892P.  
PR 22-AUG-1997; 97US-0056893P.  
PR 22-AUG-1997; 97US-0056894P.  
PR 22-AUG-1997; 97US-0056903P.  
PR 22-AUG-1997; 97US-0056908P.  
PR 22-AUG-1997; 97US-0056909P.  
PR 22-AUG-1997; 97US-0056910P.  
PR 05-SEP-1997; 97US-0056911P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057659P.  
PR 12-SEP-1997; 97US-0057761P.  
PR 02-OCT-1997; 97US-0061060P.



23-MAY-1997; 97US-0047600P.  
PR 23-MAY-1997; 97US-0047601P.  
PR 23-MAY-1997; 97US-0047612P.  
PR 23-MAY-1997; 97US-0047613P.  
PR 23-MAY-1997; 97US-0047614P.  
PR 23-MAY-1997; 97US-0047615P.  
PR 23-MAY-1997; 97US-0047616P.  
PR 23-MAY-1997; 97US-0047617P.  
PR 23-MAY-1997; 97US-0047618P.  
PR 23-MAY-1997; 97US-0047619P.  
PR 06-JUN-1997; 97US-0048964P.  
PR 06-JUN-1997; 97US-0048965P.  
PR 13-JUN-1997; 97US-0049610P.  
PR 08-JUL-1997; 97US-0051926P.  
PR 16-JUL-1997; 97US-0052874P.  
PR 18-AUG-1997; 97US-0055724P.  
PR 22-AUG-1997; 97US-0056630P.  
PR 22-AUG-1997; 97US-0056631P.  
PR 22-AUG-1997; 97US-0056632P.  
PR 22-AUG-1997; 97US-0056633P.  
PR 22-AUG-1997; 97US-0056634P.  
PR 22-AUG-1997; 97US-0056635P.  
PR 22-AUG-1997; 97US-0056636P.  
PR 22-AUG-1997; 97US-0056637P.  
PR 22-AUG-1997; 97US-0056638P.  
PR 22-AUG-1997; 97US-0056639P.  
PR 22-AUG-1997; 97US-0056640P.  
PR 22-AUG-1997; 97US-0056641P.  
PR 22-AUG-1997; 97US-0056642P.  
PR 22-AUG-1997; 97US-0056643P.  
PR 22-AUG-1997; 97US-0056644P.  
PR 22-AUG-1997; 97US-0056645P.  
PR 22-AUG-1997; 97US-0056646P.  
PR 22-AUG-1997; 97US-0056647P.  
PR 22-AUG-1997; 97US-0056648P.  
PR 22-AUG-1997; 97US-0056649P.  
PR 22-AUG-1997; 97US-0056650P.  
PR 22-AUG-1997; 97US-0056651P.  
PR 22-AUG-1997; 97US-0056652P.  
PR 22-AUG-1997; 97US-0056653P.  
PR 22-AUG-1997; 97US-0056654P.  
PR 22-AUG-1997; 97US-0056655P.  
PR 22-AUG-1997; 97US-0056656P.  
PR 22-AUG-1997; 97US-0056657P.  
PR 22-AUG-1997; 97US-0056658P.  
PR 22-AUG-1997; 97US-0056659P.  
PR 22-AUG-1997; 97US-0056660P.  
PR 22-AUG-1997; 97US-0056661P.  
PR 22-AUG-1997; 97US-0056662P.  
PR 22-AUG-1997; 97US-0056663P.  
PR 22-AUG-1997; 97US-0056664P.  
PR 22-AUG-1997; 97US-0056665P.  
PR 22-AUG-1997; 97US-0056666P.  
PR 22-AUG-1997; 97US-0056667P.  
PR 22-AUG-1997; 97US-0056668P.  
PR 22-AUG-1997; 97US-0056669P.  
PR 22-AUG-1997; 97US-0056670P.  
PR 22-AUG-1997; 97US-0056671P.  
PR 22-AUG-1997; 97US-0056672P.  
PR 22-AUG-1997; 97US-0056673P.  
PR 22-AUG-1997; 97US-0056674P.  
PR 22-AUG-1997; 97US-0056675P.  
PR 22-AUG-1997; 97US-0056676P.  
PR 22-AUG-1997; 97US-0056677P.  
PR 22-AUG-1997; 97US-0056678P.  
PR 22-AUG-1997; 97US-0056679P.  
PR 22-AUG-1997; 97US-0056680P.  
PR 22-AUG-1997; 97US-0056681P.  
PR 22-AUG-1997; 97US-0056682P.  
PR 22-AUG-1997; 97US-0056683P.  
PR 22-AUG-1997; 97US-0056684P.  
PR 22-AUG-1997; 97US-0056685P.  
PR 22-AUG-1997; 97US-0056686P.  
PR 22-AUG-1997; 97US-0056687P.  
PR 22-AUG-1997; 97US-0056688P.  
PR 22-AUG-1997; 97US-0056689P.  
PR 22-AUG-1997; 97US-0056690P.  
PR 22-AUG-1997; 97US-0056691P.  
PR 22-AUG-1997; 97US-0056692P.  
PR 22-AUG-1997; 97US-0056693P.  
PR 22-AUG-1997; 97US-0056694P.  
PR 22-AUG-1997; 97US-0056695P.  
PR 22-AUG-1997; 97US-0056696P.  
PR 22-AUG-1997; 97US-0056697P.  
PR 22-AUG-1997; 97US-0056698P.  
PR 22-AUG-1997; 97US-0056699P.  
PR 22-AUG-1997; 97US-0056700P.  
PR 05-SEP-1997; 97US-0057650P.  
PR 05-SEP-1997; 97US-0057651P.  
PR 05-SEP-1997; 97US-0057652P.  
PR 05-SEP-1997; 97US-0057653P.  
PR 05-SEP-1997; 97US-0057654P.  
PR 12-SEP-1997; 97US-0058785P.  
PR 02-OCT-1997; 97US-0061060P.  
PR 06-MAR-1998; 98WO-US0004493.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC,  
XX Bednarik DR, Endress GA, Yu G, Ni U, Feng P, Young PE, Greene JM,  
PI Perrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ender R, Brewer LA,  
PI Moore PA, Shi Y, Latleur DW, Li Y, Zeng Z, Kyaw H,  
XX  
XX WPI; 2002-634736/68.  
XX P-PSDB; ABG95227.  
XX  
XX New isolated human secreted protein for diagnosing, preventing, treating  
XX PT or ameliorating medical conditions and used as a food additive or  
XX PT preservative.  
XX  
XX Example 1; SEQ ID NO 58; 129pp; English.  
XX  
XX The invention relates to an isolated protein that is one of 186 human

CC secreted proteins, given in the specification, encoded by one of 309 cDNA  
CC sequences also given in the specification. The protein is used in a  
CC pharmaceutical composition used to prevent, treat or ameliorate a medical  
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. Disorders which are diagnosed or treated include  
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiodenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC e.g. corneal infection. The polypeptides can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,  
CC carbohydrate, vitamins, minerals, cofactors and other nutritional  
CC components. The present sequence represents a cDNA derived from a gene  
CC encoding one of the novel human secreted proteins of the invention. Note:  
CC This sequence did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPRO at  
CC [seqdata.uspro.gov/sequence.html?docID=6420526B1](http://seqdata.uspro.gov/sequence.html?docID=6420526B1)  
XX  
SQ Sequence 1049 BP; 223 A; 258 C; 303 G; 265 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 7,82e-51 Length: 1049  
Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-10-001-885-125 (1-121) x ABS73545 (1-1049)  
QY 1 MetValArgIleLeuAlaAsnGlyIleValGlnAspAspProArgValArgThr 20  
Db 116 ATGTGCGGATCTTGGCCATGGGAATGTCGACGACGACCCCGAGTGGAGAC 175  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyIleGly 40  
Db 176 ACTACCCAGCCACCAAGAGTGCATCTCCGACGAGCTTCTTCAATAGGGGCGCATGCT 235  
QY 41 AlaProProGlyIleProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAla 60  
Db 236 GCTCCCGCAGGGGGTCTGCGCCCGCCGACGACGAGGAGGCGAGCTGGGTCCTCT 295  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
Db 296 CAGTCCCGCTTCAATGACCTCAACCGGACGCTGGTGAACATGGGCTTCCGACGTGGCAT 355  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
Db 356 CTCGGCAACATGCTGTGGAGCCGGTGAACCTCCATCTCCCTCTCTCTCATGATG 415  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValThrLeuValSerHisLeuSerGln 120  
Db 416 CTGATGTTCTGGGCTCTCTCTCTGTTGGCTTGTCTTACCTGGTGGTCCACCTGATG 475  
QY 121 Arg 121  
Db 476 CGG 478  
RESULT 5  
ACD82688  
ID ACD82688 standard; cDNA; 1049 BP.  
XX  
XX ACD82688;  
XX  
XX 22-SEP-2003 (first entry)  
XX  
XX cDNA sequence #48 containing coding region of a human secreted protein.  
XX

XX	Human; secreted protein; hyperproliferative disorder; leukaemia;
KW	breast cancer; wound; reproductive disorder; blood-related disorder;
KM	haemophilia; thrombocytopaenia; immunodeficiency; thymic hypoplasia;
KM	Miskott-Aldrich syndrome; autoimmune disorder; multiple sclerosis;
KW	graft-versus-host disease; Hashimoto's thyroiditis; allergy; asthma;
KM	vital infection; bacterial infection; fungal infection; AIDS; sepsis;
KM	renal disorder; kidney failure; cardiovascular disorder; cytostatic;
KM	angina pectoris; cerebral ischaemia; congenital heart defect;
KW	respiratory disorder; neurological disorder; Alzheimer's disease;
KM	Parkinson's disease; inflammatory; Crohn's disease; vulnerability;
KM	immunosuppressive; antibacterial; haemostatic; thrombolytic;
KM	anticoagulant; neuroprotective; chemotherapeutic; antiallergic;
KM	antiaesthetic; vincristine; fungicide; anti-HIV; nephrotoxic;
KM	cerebroprotective; cardiac; neurotropic; antiparkinsonian;
XX	antiinflammatory; gene; ss.
OS	Homo sapiens.
XX	US2003049618-A1.
PN	13-MAR-2003.
XX	16-MAR-2001; 2001US-00809391.
PF	07-MAR-1997; 97US-0038621P.
XX	07-MAR-1997; 97US-0040162P.
PR	07-MAR-1997; 97US-0040163P.
PR	07-MAR-1997; 97US-0040333P.
PR	07-MAR-1997; 97US-0040334P.
PR	07-MAR-1997; 97US-0040336P.
PR	07-MAR-1997; 97US-0040626P.
PR	11-APR-1997; 97US-0043311P.
PR	11-APR-1997; 97US-0043312P.
PR	11-APR-1997; 97US-0043313P.
PR	11-APR-1997; 97US-0043314P.
PR	11-APR-1997; 97US-0043315P.
PR	11-APR-1997; 97US-0043568P.
PR	11-APR-1997; 97US-0043569P.
PR	11-APR-1997; 97US-0043570P.
PR	11-APR-1997; 97US-0043571P.
PR	11-APR-1997; 97US-0043580P.
PR	11-APR-1997; 97US-0043668P.
PR	11-APR-1997; 97US-0043670P.
PR	11-APR-1997; 97US-0043671P.
PR	11-APR-1997; 97US-0043672P.
PR	11-APR-1997; 97US-0043674P.
PR	23-MAY-1997; 97US-0047492P.
PR	23-MAY-1997; 97US-0047500P.
PR	23-MAY-1997; 97US-0047501P.
PR	23-MAY-1997; 97US-0047502P.
PR	23-MAY-1997; 97US-0047503P.
PR	23-MAY-1997; 97US-0047504P.
PR	23-MAY-1997; 97US-0047582P.
PR	23-MAY-1997; 97US-0047583P.
PR	23-MAY-1997; 97US-0047584P.
PR	23-MAY-1997; 97US-0047585P.
PR	23-MAY-1997; 97US-0047586P.
PR	23-MAY-1997; 97US-0047587P.
PR	23-MAY-1997; 97US-0047588P.
PR	23-MAY-1997; 97US-0047589P.
PR	23-MAY-1997; 97US-0047590P.
PR	23-MAY-1997; 97US-0047592P.
PR	23-MAY-1997; 97US-0047593P.
PR	23-MAY-1997; 97US-0047594P.
PR	23-MAY-1997; 97US-0047595P.
PR	23-MAY-1997; 97US-0047596P.
PR	23-MAY-1997; 97US-0047597P.
PR	23-MAY-1997; 97US-0047598P.
PR	23-MAY-1997; 97US-0047600P.
PR	23-MAY-1997; 97US-0047601P.
PR	23-MAY-1997; 97US-0047612P.
PR	23-MAY-1997; 97US-0047613P.

PR	23-MAY-1997	97US-0047614P
PR	23-MAY-1997	97US-0047615P
PR	23-MAY-1997	97US-0047617P
PR	23-MAY-1997	97US-0047618P
PR	23-MAY-1997	97US-0047632P
PR	23-MAY-1997	97US-0047633P
PR	06-JUN-1997	97US-0048664P
PR	06-JUN-1997	97US-0048670P
PR	13-JUN-1997	97US-0049410P
PR	08-JUL-1997	97US-0051926P
PR	16-JUL-1997	97US-0052874P
PR	18-AUG-1997	97US-0055724P
PR	22-AUG-1997	97US-0056630P
PR	22-AUG-1997	97US-0056631P
PR	22-AUG-1997	97US-0056632P
PR	22-AUG-1997	97US-0056645P
PR	22-AUG-1997	97US-0056662P
PR	22-AUG-1997	97US-0056664P
PR	22-AUG-1997	97US-0056676P
PR	22-AUG-1997	97US-0056877P
PR	22-AUG-1997	97US-0056878P
PR	22-AUG-1997	97US-0056880P
PR	22-AUG-1997	97US-0056881P
PR	22-AUG-1997	97US-0056882P
PR	22-AUG-1997	97US-0056884P
PR	22-AUG-1997	97US-0056886P
PR	22-AUG-1997	97US-0056887P
PR	22-AUG-1997	97US-0056888P
PR	22-AUG-1997	97US-0056892P
PR	22-AUG-1997	97US-0056893P
PR	22-AUG-1997	97US-0056894P
PR	22-AUG-1997	97US-0056903P
PR	22-AUG-1997	97US-0056908P
PR	22-AUG-1997	97US-0056939P
PR	22-AUG-1997	97US-0056911P
PR	22-AUG-1997	97US-0056912P
PR	05-SEP-1997	97US-0057650P
PR	05-SEP-1997	97US-0057659P
PR	05-SEP-1997	97US-0057761P
PR	12-SEP-1997	97US-0058785P
PR	09-OCT-1997	97US-0061660P
PR	06-MAR-1998	98WU-US0044763
PR	08-SEP-1998	98US-0014947P
PR	17-MAR-2000	2000US-0190068P
xx		
PA	(RUBE/) RUBEN S. M.	
PA	(ROSE/) ROSEN C. A.	
PA	(SOPE/) SOPPET D. R.	
PA	(CART/) CARTER K. C.	
PA	(BEDN/) BEDNARAK D. P.	
PA	(ENDR/) ENDRESS G. A.	
PA	(YUGG/) YUG G.	
PA	(NITU/) NI J.	
PA	(FENG/) FENG P.	
PA	(YOUN/) YOUNG P. E.	
PA	(GREE/) GREENE J. M.	
PA	(FERR/) FERRIE A. J.	
PA	(DUAN/) DUAN D. R.	
PA	(HUU/) HU J.	
PA	(FLOR/) FLORENCE K. A.	
PA	(OLSE/) OLSEN H. S.	
PA	(FISC/) FISCHER C. L.	
PA	(EBNE/) EBNER L.	
PA	(BREM/) BREMER L. A.	

Db 356 CTGGGACCAATCTGTGAGCGCGTGACTTCACCTCCGCTCCTTCTTCCATCATGATG 415  
|||  
QY 101 LeuGIyValArgGIyLeuLeuValGIyLeuValIlyrLeuValSerHisLeuSergIn 120  
Db 416 CTTGGTGTTCGTGGCCCTCCTCCTGGTGTGGCCCTTGTCTACCTGGTGTCACCTGAATCAG 475  
|||  
QY 121 Arg 121  
|||  
Db 476 CGG 478

RESULT 6  
ABL90057  
ID ABL90057 standard; cDNA; 1108 BP.  
XX  
XX ABL90057;  
AC  
XX  
XX ABL90057;  
DT 24-MAY-2002 (first entry)  
DE Human polynucleotide SEQ ID NO 619.  
XX  
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
XX antiallergic; hepatotropic; antidysabetic; antiinflammatory; anticancer;  
XX vulnereary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
XX neurological disease; infection; human; secreted protein; gene; ss.  
OS Homo sapiens.  
XX  
XX W0200190304-A2.  
PN 29-NOV-2001.  
ED  
PF 18-MAY-2001; 2001MO-USO16450.  
XX  
XX 19-MAY-2000; 2000US-020551SP.  
PR  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Birse CE, Rosen CA;  
PI  
DR MPI; 2002-122018/16.  
DR P-PSDB; ABB89648.  
XX  
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive,  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders.

Claim 4; SEQ ID NO 619; 2081bp + Sequence Listing; English.

The invention relates to novel genes (ABL89449-ABH90853) and proteins (ABB89040-ABB99444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 1108 BP, 234 A, 274 C, 324 G, 271 T, 0 U, 5 Other;

Alignment Scores: 8.35e-51 Length: 1108  
Aired. No.:

Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-001-885-125 (1-121) x ABL90057 (1-1108)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 147 ATGTTGGGAGATCTTGGCCATGGGGAATGTCACAGATGACACCCCGAGTGGAGACC 206  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40  
DB 207 ACTACCCAGCCACCAAGAGTAGCATCTCCGACAGAGCTTCTTCAAYAGGGCCATGCT 266  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 267 GCTCCCCCAGGGGGTCTGTGGCCCCCGCAGCAGCAGGGAGGTCACAGCTGGAGTCTGCT 326  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 327 CAGTCCCCCTTCAATGACCTCAACCGGCAGCTGTGAACATGGGCTTCCGACGTGGCAT 386  
QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuLeuPheLeuLeuMetMet 100  
DB 387 CTGGGCAACCATGCTGTGGAGCCGGTGAACCTCCATCTGCTCTTCTGCTCATGATG 446  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
DB 447 CTGGTGTGTGGCTCTCTCTCTCTGCTTGTGCTTCTACTGCTGTCTCCACCTGACTCAG 506  
QY 121 Arg 121  
DB 507 CGG 509

## RESULT 7

ABN87820 standard; cDNA; 1192 BP.

ABN87820;

12-AUG-2002 (first entry)

Human ovary specific nucleic acid SEQ ID NO:31.

Human; ovary specific nucleic acid; OSNA; ovary specific protein; OSP;  
ovary specific gene; OSG; ovarian cancer; immune response; metastasis;  
chromosome 10; gene; ss.

Homo sapiens.

MO200240535-A2.

23-MAY-2002.

20-NOV-2001; 2001WO-US045011.

20-NOV-2000; 2000US-0252061P.

27-NOV-2000; 2000US-0253257P.

(DIAD-) DIADEXUS INC.

Salceda S, Macina RA, Recipon H, Caferkey R, Sun Y, Liu C;

WPI; 2002-471617/50.

New ovary specific genes and proteins, useful as a vaccine for treating  
patients with ovarian cancer, or for diagnosing and monitoring the  
presence and metastases of ovarian cancer in a patient.

Claim 1; Page 173-174; 260pp; English.

CC AEN87790 to AEN87882 represent human ovary specific nucleic acid (OSNA)  
CC sequences, and ABB79297 to ABB79370 represent ovary specific protein  
CC (OSP) sequences from the present invention. OSNA and OSP sequences have  
CC cytostatic activity, and can be used in vaccine production and gene  
CC therapy. An antibody that specifically binds to an OSP can be used for  
CC treating a patient with ovarian cancer, particularly for inducing an  
CC immune response against the ovarian cancer cell expressing the OSNA or  
CC OSP. The OSNAs and OSPs can also be used for diagnosing and monitoring  
CC the presence and metastases of ovarian cancer in a patient

Sequence 1192 BP; 262 A; 298 C; 349 G; 283 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 9,1e-51 Length: 1192  
Score: 632.00 Matches: 121  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-001-885-125 (1-121) x AEN87820 (1-1192)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 161 ATGGTGGGAGATCTTGGCCATGGGGAATGTCGACAGACGACCCCGAGTGGAGACC 220  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHisGly 40  
DB 221 ACTACCCAGCCACCAAGAGTAGCATCTCCGACAGAGCTTCTTCAATAGGGCCATGCT 280  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 281 GCTCCCCCAGGGGGTCTGTGGCCCCCGCAGCAGCAGGGAGGTCACAGCTGGAGTCTGCT 340  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 341 CAGTCCCCCTTCAATGACCTCAACCGGCAGCTGTGAACATGGGCTTCCGACGTGGCAT 400  
QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuLeuPheLeuLeuMetMet 100  
DB 401 CTGGGCAACCATGCTGTGGAGCCGGTGAACCTCCATCTGCTCTTCTGCTCATGATG 460  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
DB 461 CTGGTGTGTGGCTCTCTCTCTCTGCTTGTGCTTCTACTGCTGTCTCCACCTGACTCAG 520  
QY 121 Arg 121  
DB 521 CGG 523

## RESULT 8

AAF82505 standard; cDNA; 550 BP.

AAF82505;

18-JUN-2001 (first entry)

Human breast tumour protein cDNA 15964.

Human; breast cancer; tumour; cytostatic; gene therapy; ss.

Homo sapiens.

MO200127276-A2.

19-APR-2001.

11-OCT-2000; 2000WO-US028255.

12-OCT-1999; 99US-00417031.

(CORI-) CORIXA CORP.



XX Harlocker SL, Dillon DC, Xu J;  
XX WPI, 2001-273773/28.  
XX  
XX New polypeptides encoded by polynucleotide sequences over-expressed in  
XX breast tumor tissue are useful to detect, monitor and treat breast  
XX cancer.  
XX  
XX Claim 13, Page 50, 52pp; English.  
XX  
XX The present sequence encodes a breast tumour-associated protein. It was  
XX shown to have at least two-fold overexpression in breast tumour tissue.  
XX The invention relates to an isolated polypeptide comprising at least an  
XX immunogenic portion of a breast tumor-specific protein, or its variant  
XX that retains the ability to react with antigen-specific antisera. The  
XX breast tumour polynucleotides, polypeptides and antibodies are useful for  
XX inhibiting development of breast cancer. The polynucleotides may be used  
XX to design primers and probes for detecting and monitoring breast cancer  
XX  
SQ Sequence 550 BP; 163 A; 163 C; 164 G; 118 T; 0 U; 3 Other;  
SO  
Alignment Scores:  
Pred. No.: 6.13e-46 Length: 550  
Score: 577.00 Matches: 117  
Percent Similarity: 97.52% Conservative: 1  
Best Local Similarity: 96.69% Mismatches: 2  
Query Match: 91.30% Indels: 2  
DB: Gaps: 0  
US-10-001-885-125 (1-121) x AAF82505 (1-550)  
QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspAspProArgValArgThr 20  
DB 115 ATGGTGGGATCTTGGCCATGGGAAATGCTGCAGACGACCCCGAGTGAAGACC 174  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPheAsnArgGlyHISGly 40  
DB 175 ACTACCCAGCCACCAAGAGTACGATTCCTGCAGAGCTTCTTCATAGGGGCGCATGCT 234  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 235 GCTCCGCCAGGGGCTCTGGCCCGCCGACAGCAGGAGGCTGCGAGGCTGCT 234  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTrpHis 80  
DB 295 CAGTCCCGCTTCATGACCTCAACCGGACGCTGTGAACATGGGCTTCCGACGTGGCAT 354  
QY 81 LeuGlyAsnHisAla-ValGlnProValThrSerIleLeuLeuPheLeuLeuMetMe 100  
DB 355 CTCGGCAACCATGCTTGTGAGCGCGGTGACCTCATCTGCT-CTTCTCGTCAAGAT 413  
QY 100 tLeuGlyValArgGlyLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerG1 120  
DB 414 GCTTGGTGGTGGTGGCTTCTGCTGTGGCTGTGCTACTGCTGTGCCACCTGATCA 473  
QY 120 n 120  
DB 474 A 474  
RESULT 9  
AA64820 ID AA64820 standard; cDNA; 1165 BP.  
XX AC AA64820;  
XX 13-FEB-2002 (first entry)  
XX DNA encoding novel human diagnostic protein #624.  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX

OS Homo sapiens.  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX WPI, 2001-639362/73.  
XX P-PDB; ABG00633.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
XX Claim 1, SEQ ID NO 624; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping.  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. AA64197-AA94564 represent novel human diagnostic  
XX coding sequences of the invention. Note: The sequence data for this  
XX patent did not appear in the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pat\_sequences  
XX  
SQ Sequence 1165 BP; 247 A; 279 C; 342 G; 297 T; 0 U; 0 Other;  
SO  
Alignment Scores:  
Pred. No.: 5.79e-32 Length: 1165  
Score: 434.00 Matches: 116  
Percent Similarity: 86.67% Conservative: 1  
Best Local Similarity: 85.93% Mismatches: 3  
Query Match: 68.67% Indels: 15  
DB: Gaps: 3  
US-10-001-885-125 (1-121) x AA64820 (1-1165)  
QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspAspProArgValArgThr 20  
DB 352 ATGGTGGGATCTTGGCCATGGGAAATGCTGCAGACGACCCCGAGTGAAGACC 411  
QY 20 rThrThrGlnProProArgGlySerIleProArgGlnSer-PhePheAsnArgGlyHISG 40  
DB 412 CACTACCCAGCCACCAAGAGTACGATTCCTGCAGAGCTTCTTCATAGGGGCGCATG 471  
QY 40 lYAlaProProGlyGlyPro-GlyProArgGlnGlnAlaGly-----AlaArgLeu 57  
DB 472 GTGCTCCCGCCAGGGGCTCGTGGCCCGCCG---CAGCAAGCAGGAGGCTGCAAGGCTG 528  
QY 58 GlyAlaAlaGln-SerProPheAsnAsp-LeuAsnArgGlnLeuVal-AsnMetGlyPhe 76

DB 529 GGTGCTGCTAGTCCCTCCCTTCAATGACCTTAAACCGGACGCTGCTGAACATGGGCTTT 588  
 QY 77 Progin-TrypHisLeu-GlyAsnHisAlaVal---GluProVal-ThrSerIleLeu-Le 94  
 DB 589 CCGGAGTTGGAGCTCTTGGCAACCATGCTGTGGGAGCCGGTAAGATCATCTCTCCCT 648  
 QY 94 ULeuPheLeuLeuMetLeuGlyValArgGlyLeuLeuValGlyLeuValIle 114  
 DB 649 CCTCTTCCTGCTCATGATGCTGTGGTCTGGCCCTCCTCTGTGGCCCTTGTACTCT 708  
 QY 114 UValSerHisLeu-SerGlnArg 121  
 DB 709 GTGTCCCACTGAGTCAAGCGG 731

RESULT 10  
 ABX74673/C  
 ID ABX74673 standard; cDNA; 517 BP.

AC ABX74673;  
 DT 21-MAR-2003 (first entry)

DE Human cDNA sequence #128 differentially expressed in CC-RCC types.

KM Human; microarray; solid surface; immobilised probe; CC-RCC;  
 KM differential expression profile; aggressive CC-RCC tumour type;  
 KM non-aggressive CC-RCC tumour type; clear cell renal carcinoma;  
 KM gene expression profiling; tumour tissue; gene; se.

OS Homo sapiens.

PN MO200279411-A2.

PD 10-OCT-2002.

PF 29-MAR-2002; 2002WO-US009576.

PR 29-MAR-2001; 2001US-0279411P.

PA (VAND-) VAN ANDEL INST.

PI Haab B, Rhodes D, Teh BT, Takashi M;

PI WPI; 2003-040679/03.

PT New microarray, comprising a matrix of cDNA probe from a set of probes  
 PT immobilized to a solid surface in predetermined order, useful in the  
 PT prognosis of patients with clear cell renal carcinoma.

PS Claim 1; SEQ ID NO 459; 179bp; English.

XX The present invention relates to a microarray comprising a matrix of at  
 CC least one cDNA probe from a set of probes immobilised to a solid surface  
 CC in a predetermined order, where a row of pixels corresponds to replicates  
 CC of one distinct probe from the set. The probes are complementary to  
 CC nucleic acid sequences that are expressed differentially in aggressive as  
 CC compared to non-aggressive types of clear cell renal carcinoma (CC-RCC)  
 CC and that hybridise to the probes under high stringency conditions. The  
 CC microarray is useful for the prognosis of patients with CC-RCC, wherein  
 CC aggressive and non-aggressive CC-RCC tumour types are characterised by  
 CC differential expression profiles of genes that hybridise with one or more  
 CC probes immobilised on the microarray. The arrays are useful for gene  
 CC expression profiling of tumour and normal tissues. The present sequence  
 CC represents a human cDNA sequence differentially expressed in aggressive  
 CC vs. non-aggressive type CC-RCC phenotypes

XX Sequence 517 BP; 166 A; 154 C; 90 G; 107 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.0705 Length: 517  
 Score: 113.00 Matches: 29  
 Percent Similarity: 96.67% Conservative: 0

Query Match: 17.88% Indels: 1  
 DB: 7 Gaps: 0  
 US-10-001-885-125 (1-121) x ABX74673 (1-517)

QY 92 IleLeuLeuPheLeuLeuMetLeuGlyValArgGlyLeuLeuValGlyLeu 111  
 DB 515 ATCTGCTCTCTCTCTCTGCTCATGATGCTGTGGTCTGGCCCTCCTCTGTGGCTT 456

QY 112 ValIleuValSerHisLeuSerGlnArg 121  
 DB 455 GTCTACTCT-GTGTCCCACTGAGTCAAGCGG 427

RESULT 11  
 ADE07153  
 ID ADE07153 standard; DNA; 1110 BP.

AC ADE07153;

DT 29-JUN-2004 (first entry)

DE Novel coding sequence (useful for identifying genetic disorders) #219.

KM novel gene; novel protein; tissue marker; molecular weight marker;

KM chromosome marker; genetic disorder; gene; ds.

OS Unidentified.

PN MO2003054152-A2.

PN 03-JUL-2003.

PF 10-DEC-2002; 2002WO-US039555.

PR 10-DEC-2001; 2001US-039739P.

PR 11-DEC-2001; 2001US-0339453P.

PR 14-MAR-2002; 2002US-0365091P.

PR 14-MAR-2002; 2002US-0365384P.

PR 12-APR-2002; 2002US-0372381P.

PR 12-APR-2002; 2002US-0372615P.

PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao Q, Wang J;

PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;

PI WPI; 2003-569235/53.

PI P-PsDB; ADE08064.

PT New polynucleotides, useful for expressing recombinant proteins for

PT analysis, characterization or therapeutic use, or as markers for tissues

PT in which the corresponding protein is preferentially expressed.

PS Claim 1; SEQ ID NO 219; 1177bp; English.

XX The invention comprises the amino acid and coding sequences of novel

CC proteins. The DNA and protein sequences of the invention are useful as:

CC markers for tissues in which the corresponding protein is preferentially

CC expressed; as molecular weight markers on gels; as chromosome markers or

CC tags; to identify chromosomes or to map related gene positions; and to

CC compare with endogenous DNA sequences in patients to identify potential

CC genetic disorders. The present DNA sequence represents a gene of the

CC invention.

XX Sequence 1110 BP; 330 A; 243 C; 344 G; 193 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.302 Length: 1110  
 Score: 110.50 Matches: 30  
 Percent Similarity: 50.00% Conservative: 17

XX	WPI; 2003-058429/05.
DR	P-PsDB; ABB84669.
PT	Novel human secreted protein useful for treating, preventing or diagnosing cancer, hepatitis, psoriasis, actinia, diabetes mellitus, anemia, epilepsy, cataract, Alzheimer's disease.
PS	Claim 101; Page 185-186; 188pp: English.
CC	This invention describes novel secreted proteins (SPCP) which have antiarteriosclerotic, antithrombotic, hepatotropic, cytotatic, anti-HIV, anti-allergic, antidiabetic, antianemic, antidiabetic, anti-inflammatory, neuroprotective, anticancer, antiproliferative, vasotrophic, antihypertensive, cardiostimulant, hypotensive, anticonvulsant, neurotropic, immunosuppressive, antiparkinsonian and ophthalmological activity. The polypeptides and polypeptides of the invention can be used for diagnosing, treating or preventing cell proliferative disorder e.g., arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, cancer, autoimmune/inflammatory disorders e.g., acquired immunodeficiency syndrome (AIDS), allergies, asthma, anaemia, diabetes mellitus, Crohn's disease, multiple sclerosis, ulcerative colitis, psoriasis, rheumatoid arthritis, etc.; cardiovascular disorder e.g., myocardial infarction, angina pectoris, hypertension, Raynaud's disease, myocarditis, pericarditis, etc.; neurological disorders e.g., epilepsy, Huntington's disease, Parkinson's disease, Alzheimer's disease, Creutzfeldt-Jakob disease, etc. and developmental disorders e.g., Duchenne and Becker muscular dystrophy, cataract, gonadal dysgenesis, Cushing's syndrome, etc. The products of the invention can also be used for drug screening, proteome analysis, microarrays creating knock-in humanised animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals, and as hybridization probes for mapping naturally occurring genomic sequences. ABB57545-ABB57569 represent cDNA's encoding the secreted proteins represented by ABB84649-ABB84673, described in the disclosure of the invention
XX	
QO	Sequence 2999 BP, 912 A; 554 C; 585 G; 948 T; 0 U; 0 Other;
AS	Alignment Scores:
Pred. No.:	0.981 Length: 2999
Score:	110.50 Matches: 30
Percent Similarity:	50.00% Conserved: 17
Best Local Similarity:	31.91% Mismatches: 34
Query Match:	17.48% Indels: 13
DB:	Gaps: 3
US-10-001-885-125 (1-121) x ABB57565 (1-2999)	
OY	26 ARGGLYSERILPEPRARGINSEK-----PHEPHENARGLYLHISGLYALA 41
DB	547 CGGGACAGCGGCCGAGAAGAGCAGCAAGATTTGAAGCTTACAGAACCAACTGGT 606
OY	42 PROPGILGLIPROGLIPROAGINGLNGINALGLYLAARGLEUGLYLAALAGIN 61
DB	607 GACCCTGTGGAGAGACACACAG-----AAATGGGAACAATT--- 645
OY	62 SERPROPHASNAPLEUNSNRGLILEUVALAMMETGLYPHEPROGINTPRHSLEU 81
DB	646 -----TTTGTCGTAAGTGAACAAAACCTTTTCACACATGGCTTCACAGATGTATTT 659
OY	82 GLVASHNIALVALGLIPROVALTHRSERILEUENLEUPHELEUMMETLEU 101
DB	700 GGAGAACGAATAGTGGACCACTAATATCATTTCTTTGGGTATGCCTGTGCTT 759
OY	102 GLVVALRGYLEULEUENLEUVALGLYLEUVALTYRLEUVAL 115
DB	760 GGCTGTCAAGCCCTTGACACTAGTCTGCTCTTTGGCCCTGTT 801
RESULT 13	
ACD03995	
ID	ACA03995 standard; cDNA; 2668 BP.

XX ACA03995;  
AC  
XX  
XX 27-MAY-2003 (first entry)  
DT  
XX  
XX  
DE cDNA downregulated in senescent cells Incyte ID NO: 403703.1.

Human; senescence; ss; gene; cancer; proliferative disorder; leukaemia;  
KM adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; teratocarcinoma;  
KM adrenal gland cancer; bladder cancer; bone cancer; bone marrow cancer;  
KM brain cancer; breast cancer; cervical cancer; colon cancer; heart cancer;  
KM oesophageal cancer; gall bladder cancer; ganglial cancer; kidney cancer;  
KM liver cancer; lung cancer; muscle cancer; ovarian cancer; penile cancer;  
KM pancreatic cancer; parathyroid gland cancer; prostate cancer;  
KM salivary gland cancer; skin cancer; small intestine cancer;  
KM spleen cancer; stomach cancer; testicular cancer; thyroid cancer;  
KM thyroid cancer; uterine cancer.

XX Homo sapiens.  
XX  
XX US2002192678-A1.  
XX  
XX 19-DEC-2002.  
XX  
XX 07-FEB-2002; 2002US-00071766.  
XX  
XX 09-FEB-2001; 2001US-0268380P.  
XX  
XX (CHEN/) CHEN H.  
XX  
XX Chen H;  
XX  
XX WPI: 2003-328858/31.  
XX  
XX

PT New combination comprising cDNAs or their complements, useful for  
PT detecting changes in expression of genes encoding proteins associated  
PT with senescence, and in diagnosing, staging or treating proliferative  
PT diseases, e.g. cancer.  
XX  
XX

PS Claim 9; Page 143-144; 195pp; English.

XX The invention relates to a combination comprising a plurality of cDNAs,  
CC or their complements that are differentially expressed in cancer and  
CC other proliferative disorders. The combination is useful in detecting  
CC changes in expression of genes encoding proteins that are associated with  
CC senescence or in diagnosing, staging, treating, or monitoring the  
CC progression or treatment of subjects with proliferative diseases such as  
CC cancer e.g. adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma,  
CC sarcoma, teratocarcinoma; cancer of the adrenal gland, bladder, bone,  
CC bone marrow, brain, breast, cervix, colon, oesophagus, gall bladder,  
CC ganglial, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid  
CC gland, penis, prostate, salivary glands, skin, small intestine, spleen,  
CC stomach, testis, thymus, thyroid and uterus. The present sequence  
CC represents cDNA of genes that are downregulated in senescent cells  
XX  
XX

XX Sequence 2668 BP; 893 A; 407 C; 464 G; 903 T; 0 U; 1 Other;

XX Alignment Scores:

Score: 1.84 Length: 2668  
Pred. No.: 107.00 Matches: 22  
Percent Similarity: 61.54% Conservative: 10  
Best Local Similarity: 42.31% Mismatches: 20  
Query Match: 16.93% Gaps: 0  
DB: 7

US-10-001-885-125 (1-121) x ACA03995 (1-2668)

QY 64 PheAaApLeuAaArgGlnLeuValAsmMetGlyPheProGlnTrpHisLeuGlyAsn 83  
DB 347 TTGGTGAACCTGAACAAACCTTATCAACAGGGCTTCACAGAGATGATTTGGAGAA 406  
QY 84 HisAlaValGluProValThrSerIleLeuLeuLeuPheLeuLeuMetMetLeuGlyVal 103

DB 407 CGAATAGTGCAGACCACTAATAGCATTTCTTTGGCTTATGCTGTGTTCTTGGCCG 466  
QY 104 ArgGlyLeuLeuValGlyLeuValTyrLeuVal 115  
DB 467 CAAGCCCTTGACCTAGTGTCTGTCCTTTCGCTTGT 502

RESULT 14

AAS45077/C  
ID AAS45077 standard; cDNA, 4177 BP.

AC AAS45077;  
XX

DT 18-DEC-2001 (first entry)  
XX

DE cDNA encoding novel human secretory protein, Seq ID No 158.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;  
KM ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;  
KM transgenic animal; Alzheimer's disease; Parkinson's disease; burn;  
KM amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;  
KM ulcer; osteoporosis; bone degenerative disorder; periodontal disease;  
KM gut protection; lung; liver fibrosis; immune deficiency; infection;  
KM severe combined immunodeficiency; SCID; autoimmune disorder; allergy;  
KM multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;  
KM fertility; analgesic; pain; antigen; ss.

XX Homo sapiens.  
XX  
XX

XX W0200166689-A2.  
XX  
XX 13-SEP-2001.  
XX  
XX 05-MAR-2001; 2001WO-US04942.  
XX  
XX

XX 07-MAR-2000; 2000US-00519705.  
XX 19-MAY-2000; 2000US-00574454.  
XX 17-JUN-2000; 2000US-00596193.  
XX 14-JUL-2000; 2000US-00616847.  
XX 19-SEP-2000; 2000US-00663363.  
XX 20-OCT-2000; 2000US-00693267.  
XX  
XX

XX (HYSE-) HYSEQ INC.  
XX

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;  
PI Zhao QA, Yang Y, Dzmanac RT, Zhang J, Chen R, Xue AJ, Wang J;  
XX WPI: 2001-589934/66.  
XX P-PSDB; AAU28177.  
XX

PT Novel polypeptides and nucleic acids obtained from cDNA libraries  
PT prepared from various human tissues, for diagnosis and treatment of  
PT cancer, neurological, inflammatory, and autoimmune disorders.  
XX  
XX

PS Claim 1; SEQ ID NO 158; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)  
CC and polynucleotides (II). (I) and (II) are useful for treating  
CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,  
CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is  
CC involved in increasing haematopoiesis, stem cell survival, bone growth  
CC and remodeling (I), (II) and modulators of (II) are useful for  
CC prophylaxis or treatment of one or more cancers. (II) is also useful for  
CC creating transgenic animals useful for studying the in vivo activities of  
CC the polypeptide as well as for studying modulators of the polypeptides.  
CC (I) induces the proliferation of neural cells and regeneration of nerve  
CC and brain tissue and is useful for the treatment of central and  
CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,  
CC Parkinson's disease, Huntington's disease, and amyotrophic lateral  
CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic  
CC activity, regulation of haematopoiesis and is useful for treating myeloid  
CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia  
CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve  
CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,



Tue Aug 24 09:48:27 2004

us-10-001-885-125.p2n.rng

Page 14

Db : |||||  
3689 TAGCCCTTGGGTTCCCTGGAGCCCTTCTC 3717

Search completed: August 22, 2004, 08:37:48  
Job time : 347 secs

CM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 22, 2004, 07:19:21 ; Search time 2464 Seconds

(without alignments)  
1466.446 Million cell updates/sec

Title: US-10-001-885-125

Perfect score: 632  
Sequence: 1 MVRILANGELIVQDDPRVRT.....GVKGLLVGLVYLVSLSQR 121

Scoring table:

BLOSUM62  
Xgapext 10.0, Xgapext 0.5  
Ygapext 10.0, Ygapext 0.5  
Rgapext 6.0, Rgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seg length: 0  
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=xlp  
-Q=/cgn2\_1/USPRO.spool/p/US1001885/runat.17082004.151718.25914/app.query.fasta\_1.263  
-DB=EST -QMT=fastap -SUFFIX=p2n.txt -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US1001885.GCGN.1.1.5180 @runat.17082004.151718.25914 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_piro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrt:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	632	100.0	505	14	CA942977 1166F05.Y
2	632	100.0	566	14	CB926571 12B2027
3	632	100.0	688	12	BI521387 603081177
4	632	100.0	713	12	BI193763 602948418
5	632	100.0	714	10	BE568253 602184629
6	632	100.0	765	13	BK095713 BK095713
7	632	100.0	980	10	BE904302 BK095713
8	632	100.0	1104	12	BC331398 602433369
9	618	97.8	528	10	AM960624 EST372655
10	602	95.3	1059	13	BK420823 BX420823
11	592	93.7	447	9	AA452397 7X29404.Y
12	574.5	90.9	454	10	AM462809 BP230010B
13	574.5	90.9	468	10	BE237464 BE237464
14	574.5	90.9	523	10	BE664644 152244.MA
15	574.5	90.9	537	10	BF046179 BP250022B
16	574.5	90.9	569	10	BE588875 194533.BA
17	574.5	90.9	694	14	CB437353 CB437353
18	570.5	90.3	527	9	AV591976 AV591976
19	570.5	90.3	653	13	BU671222 NISC-1706
20	568.5	90.0	600	14	CA527226 8035-48.M
21	568.5	90.0	651	10	BE628297 BE628297
22	568.5	90.0	783	14	CK129164 AGENCOURT
23	568.5	90.0	784	12	BI658578 602384311
24	568.5	90.0	802	12	BG662915 602797603
25	568.5	90.0	823	10	BR161467 BR161467
26	568.5	90.0	891	12	BI905000 603169151
27	568.5	90.0	949	13	BU511854 AGENCOURT
28	568.5	90.0	1095	11	AK008338 Mus muscu
29	568.5	90.0	1732	11	AK035688 Mus muscu
30	567.5	89.8	546	14	CP363802 833705.MA
31	566.5	89.6	658	14	CP181245 B18054.MA
32	564.5	89.3	553	13	BU088757 BU088757
33	564.5	89.3	558	10	BE074572 SNEST4A89
34	562.5	89.0	526	12	EX523683 BX523683
35	542.5	85.8	519	29	CG607245 OST285976
36	542	85.8	631	12	BM714109 UI-E-ETO-
37	540.5	85.5	691	13	BY708262 BY708262
38	539.5	85.4	1069	12	BI414577 BI414577
39	537.5	85.0	443	9	AA428101 7W49H02.Y
40	537.5	85.0	504	29	CG607398 OST186267
41	529.5	83.8	350	10	BE237465 146772.MA
42	524	82.9	1027	12	BI522220 603081177
43	523.5	82.8	474	9	AA032373 m14406.Y
44	513.5	81.2	497	29	CG648434 OST40197
45	507.5	80.3	505	9	AA271091 va71b04.Y

## ALIGNMENTS

RESULT 1  
CA942977  
LOCUS CA942977 505 bp mRNA linear EST 30-DEC-2002  
DEFINITION t166f05.y1 HR85 islet Homo sapiens CDNA clone IMAGE:607449 5',  
CA942977  
ACCESSION CA942977 GI:27431457  
VERSION CA942977.1 GI:27431457  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
1 (bases 1 to 505)

## AUTHORS

Melton, D., Brown, J., Keny, G., Permut, A., Lee, C., Kaestner, K., Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Birstein, A., Schmidt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, W., Williams, T., Jackson, I., and Bowers, Y.

## TITLE

## JOURNAL

Unpublished (2000)  
Endocrine Pancreas Consortium

## COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Harvard University, Howard Hughes Medical Institute  
Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
Tel.: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@mol.biol.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@mol.wustl.edu)  
Seq primer: -400P from Glibco

High quality sequence stop: 463.

## FEATURES

## source

Location/Qualifiers  
1..505

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6607449"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_lib="HR85 islet"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
NotI; Site\_2: XhoI; cDNA made by oligo-dT priming;  
Size selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permut Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110. E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

## ORIGIN

## Alignment Scores:

Pred. No.:	1..948-45	Length:	505
Score:	632.00	Matches:	121
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-10-001-885-125 (1-121) x CA942977 (1-505)

QY 1 MetValArgGileuAlaAsnGlyGluIleValAGlnAspAspProArgValArgThr 20  
DB 123 ATGGTGGGATTTGGCCATGGGAAATCGTGCAGATGACGACCCCGGATGAGACC 182  
QY 21 ThrThGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 183 ACTACCGACGACCAAGAGTGCATTCCTCGACAGACTCTTCAACAGGGCCATGCT 242  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 243 GCTCCCGGAGGGGCTCTGGCCCGCCGACGACGAGGAGGCGGAGGCTGGGCTCT 302  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTyrHis 80  
DB 303 CAGTCCCGCTTCAATGACTCAACCGGACGCTGGGAAATGGGCTTCCGACATGGCAT 362  
QY 81 LeuGlyAsnHisAlaValGlnProValThrSerIleLeuLeuLeuPheLeuLeuMetMet 100  
DB 363 CTGGGCAACCAAGCTGTGAGCGCGGAGCTTCATCTGCTCTCTCTCTCTCTCTCTCT 422

## DB

423 CTGGATTTCTGGGCTCTCTCTCTGTTGGCTTGTCTACCTGGTGGTCACTAGTACG 482

## QY

121 Arg 121

## DB

483 CCG 485

## RESULT 2

## LOCUS

## DEFINITION

CB296571 566 bp mRNA linear EST 28-FEB-2003  
12B22027\_rev\_1\_B02\_r\_025.ab1 Chimpanzee brain library Koo's Pan  
troglodytes cDNA clone 12B22027\_rev\_1\_B02\_r\_025.ab1 5', mRNA  
sequence.

## ACCESSION

CB296571 GI:28622001

## KEYWORDS

EST.

## SOURCE

Pan troglodytes (chimpanzee)

## ORGANISM

Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

## REFERENCE

1 (bases 1 to 566)  
Hellmann, I., Zolner, S., Enard, W., Ebersberger, I., Nickel, B. and

Paabo, S.  
Selection on human genes as revealed by comparisons to chimpanzee

## AUTHORS

Genome Res. (2003) In press.

## TITLE

Evolutionary Genetics

## JOURNAL

Evolutionary Genetics

## COMMENT

Max Planck Institute for evolutionary Anthropology  
Deutscher Platz 6, 04103 Leipzig, Germany  
Tel: +49-(0)-341-3550 500  
Fax: +49-(0)-341-3550 555  
Email: paabo@eva.mpg.de  
Seq primer: M13 reverse.

## FEATURES

Location/Qualifiers  
1..566

/organism="Pan troglodytes"  
/mol\_type="mRNA"  
/db\_xref="taxon:9598"  
/clone="12B22027\_rev\_1\_B02\_r\_025.ab1"  
/sex="male"  
/tissue\_type="Brain, presumably cortex"  
/dev\_stage="adult"  
/lab\_host="Epizurian Coli (TM) XL-10-Gold"  
/clone\_lib="Chimpanzee Brain library Koo's"  
/note="Vector: pUC19; Site\_1: SfiI-B; The  
library was prepared using the SMART cDNA library  
construction kit (Clontech), doing only primer extension,  
but not PCR amplification of the cDNA. The only deviation  
from the published protocol was that we cloned the cDNA  
into a plasmid vector."

## ORIGIN

## Alignment Scores:

Pred. No.:	2..286-45	Length:	566
Score:	632.00	Matches:	121
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-10-001-885-125 (1-121) x CB296571 (1-566)

QY 1 MetValArgGileuAlaAsnGlyGluIleValAGlnAspAspProArgValArgThr 20  
DB 169 ATGGTGGGATTTGGCCATGGGAAATCGTGCAGATGACGACCCCGGATGAGACC 228  
QY 21 ThrThGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
DB 229 ACTACCGACGACCAAGAGTGCATTCCTCGGACAGACTCTTCAACAGGGCCATGCT 288





ORIGIN Note: this is a NIH\_MGC Library. | "

Alignment Scores: 3.14e-45 Length: 713  
Pred. No.: 632.00 Matches: 121  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 12 Gaps: 0

US-10-001-885-125 (1-121) x B1193763 (1-713)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 31 ATGTGGCGATCTTGGCCATATGGGAATCGTGCAGATGACGCCCGGAGTGGAGCC 90  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyValIleGly 40  
DB 91 ACTACCCAGCCACCAAGAGTGAAGCTTCTCGACAGAGCTTCTTCAATAGGGGCCATGGT 150  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 151 GGTCCCGCAGGGGGGCTCTGGCCCCCGCAGCAGCAGGAGTGGCCAGTGGTGGTGGT 210  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnValGlyPheProGlnTrpHis 80  
DB 211 CAGTCCCCCTTCAATGACCTCAACCGGCGAGCTGTGAACATGGGCTTCCGCACTGGCAT 270  
QY 81 LeuGlyAsnHisAlaValAlaGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
DB 271 CTGGGCAACCATGCTGTGGAGCCGGTGAACCTCCATCTCTCTCTCTCTCTCTCTCTAT 330  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValIleValIleValSerHisLeuSerGln 120  
DB 331 CTGGGTGTGGGCT 390  
QY 121 Arg 121  
DB 391 CGG 393

RESULT 5  
BF568253 714 bp mRNA linear EST 12-DEC-2000  
LOCUS 602184629F1 NIH\_MGC\_42 Homo sapiens cDNA clone IMAGE:4300518 5',  
DEFINITION mRNA sequence.  
BF568253  
ACCESSION BF568253.1 GI:11641633  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 714)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov

FEATURES  
Plate: L10C1159 row: j column: 07  
High quality sequence stop: 712.  
Location/Qualifiers  
1..714  
/organism="Homo sapiens"

/clone="IMAGE:4300518"  
/cissue type="epithelioid carcinoma cell line"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH MGC 42"  
/note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT printing;  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAAGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library. | "

ORIGIN  
Alignment Scores: 3.14e-45 Length: 714  
Pred. No.: 632.00 Matches: 121  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 10 Gaps: 0

US-10-001-885-125 (1-121) x BF568253 (1-714)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
DB 153 ATGTGGCGATCTTGGCCATATGGGAATCGTGCAGATGACGCCCGGAGTGGAGCC 212  
QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyValIleGly 40  
DB 213 ACTACCCAGCCACCAAGAGTGAAGCTTCTCGACAGAGCTTCTTCAATAGGGGCCATGGT 272  
QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
DB 273 GGTCCCGCAGGGGGGCTCTGGCCCCCGCAGCAGCAGGAGTGGCCAGTGGTGGTGGT 332  
QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnValGlyPheProGlnTrpHis 80  
DB 333 CAGTCCCCCTTCAATGACCTCAACCGGCGAGCTGTGAACATGGGCTTCCGCACTGGCAT 392  
QY 81 LeuGlyAsnHisAlaValAlaGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
DB 393 CTGGGCAACCATGCTGTGGAGCCGGTGAACCTCCATCTCTCTCTCTCTCTCTCTCTAT 452  
QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValIleValIleValSerHisLeuSerGln 120  
DB 453 CTGGGTGTGGGCT 512  
QY 121 Arg 121  
DB 513 CGG 515

RESULT 6  
BX095713 765 bp mRNA linear EST 04-FEB-2003  
LOCUS BX095713 Soares fetal N2HFE\_9w Homo sapiens cDNA clone  
DEFINITION IMAGE998B081942; IMAGE:787879, mRNA sequence.  
BX095713  
ACCESSION BX095713.1 GI:27842493  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 765)  
AUTHORS Ebert L., Heil O., Hennig S., Neubert P., Partsch B., Peters M.,  
Radelof U., Schneider D. and Korn B.  
TITLE Human Unigenesec - RZPD  
JOURNAL Unpublished (2003)  
COMMENT Contact: Ina Rolfs

RZPD: IMAGE998B081942.  
 RZPDLIB: I.M.A.G.E./CDNA Clone Collection;  
 Human Unigeneset - RZPD3 (RZPDLIB No.972)  
<http://www.rzpd.de/Cloncards/cgi-bin/showlib.pl.cgi?response=libNo=972> Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de  
 This clone is available royalty-free from RZPD;  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 M13r, Primer sequence: TTTCACACAGAAACAGCATATAC.  
 Location/Qualifiers

## FEATURES

Source

1..765  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE998B081942 ; IMAGE.787879"  
 /dev\_stage="8-9 weeks"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares, total\_fetus\_Nb2H8\_9w"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from pooled 8-9 week  
 (total) fetus material with a Not I - oligo(dt) primer [5'  
 TGTTACCACTGACGAGGAGGCGGCGTTATTTTCTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

## Alignment Scores:

Pred. No.: 3,46e-45 Length: 765  
 Score: 632.00 Matches: 121  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-001-885-125 (1-121) x BX095713 (1-765)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
 Db 45 ATGGTGGGAGATCTTGGCCAAATGGGGAATCGTGACAGATGACGCCCGAGTGAAGACC 104  
 QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
 Db 105 ACTACCCAGCCACCAAGAGGTAGCATTCCTGACAGAGCTTCTTCAACAGGGGCCATGGT 164  
 QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
 Db 165 GCTCCCCAGGGGGTCTCTGGCCCCCGCCAGCAGCAGGCGAGGTGCGAGGCTGGGCTGCT 224  
 QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTTPHis 80  
 Db 225 CAGTCCCCCTTCATATGACCTCAACCGGACAGCTGTGAACATGGGCTTCCGAGTGGCAT 284  
 QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
 Db 285 CTTCGCAACCAATGCTGTGAGCCGCTGACCTCACTGCTCTTCTTCTGCTCATATG 344  
 QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120  
 Db 345 CTTCGTTGTTGCTGGCTCTCTCTGTTGGCTTGTCTAAGCTGCTGCCACCTGAGTCAG 404  
 QY 121 Arg 121  
 Db 405 CGG 407

RESULT 7

BE904302  
 LOCUS BE904302 980 bp mRNA linear EST 20-OCT-2000  
 DEFINITION 601494571P2 NIH\_MGC\_70 Homo sapiens cDNA clone IMAGE.3896656 5',  
 mRNA sequence.  
 ACCESSION BE904302  
 VERSION BE904302.1 GI:10396415  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 980)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgs@bbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LHM9689 row: n column: 17  
 High quality sequence stop: 711.  
 Location/Qualifiers

## FEATURES

source

1..980  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3896656"  
 /tissue\_type="epithelioid carcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 70"  
 /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.1 kb. Library constructed by Life  
 Technologies."

## ORIGIN

## Alignment Scores:

Pred. No.: 4,88e-45 Length: 980  
 Score: 632.00 Matches: 121  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-001-885-125 (1-121) x BE904302 (1-980)

QY 1 MetValArgIleLeuAlaAsnGlyGluIleValGlnAspAspProArgValArgThr 20  
 Db 140 ATGGTGGGAGATCTTGGCCAAATGGGGAATCGTGACAGATGACGCCCGAGTGAAGACC 199  
 QY 21 ThrThrGlnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisGly 40  
 Db 200 ACTACCCAGCCACCAAGAGGTAGCATTCCTGACAGAGCTTCTTCAATGGGGCCATGGT 259  
 QY 41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla 60  
 Db 260 GCTCCCCAGGGGGTCTCTGGCCCCCGCCAGCAGCAGGCGAGGTGCGAGGCTGGGCTGCT 319  
 QY 61 GlnSerProPheAsnAspLeuAsnArgGlnLeuValAsnMetGlyPheProGlnTTPHis 80  
 Db 320 CAGTCCCCCTTCATATGACCTCAACCGGACAGCTGTGAACATGGGCTTCCGAGTGGCAT 379  
 QY 81 LeuGlyAsnHisAlaValGluProValThrSerIleLeuLeuPheLeuLeuMetMet 100  
 Db 380 CTTCGCAACCAATGCTGTGAGCCGCTGACCTCACTGCTCTTCTGCTCATATG 439  
 QY 101 LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrLeuValSerHisLeuSerGln 120

**Tue Aug 24 09:48:29 2004**

us-10-001-885-125.p2n.rst

Page 6

Db	440	CTTGAGTTCGAGGACCTCCTCGGTGGCTGTGACCTGGATGCCACCTGATGAC	499
Oy	121		
Db	500	CGG 502	
RESULT 8			
BG331398			
LOCUS			
DEFINITION			
ACCESSION	BG331398	1104 bp	mRNA
VERSION	6024433369F1		linear
KEYWORDS	NRH_MGC_18	Homio sapiens	cdNA clone IMAGE:4550896 5',
SOURCE	BG331398		
ORGANISM	EST.		
	Homio sapiens		
	Homio sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 1104)		
JOURNAL	NRH-MGC http://mgc.nci.nih.gov/.		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		

Tissue Procurement : DCTD/DP/Gazdar  
 cDNA Library Preparation : Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by : The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by : Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 plate: LLCM1244 row: j column: 17  
 High quality sequence stop: 636.  
 Location/Qualifiers  
     1..1104  
FEATURES  
     source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4550896"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH_MGC_18"
/notes="Organ: Lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: This is a
NH_MGC Library."

```

Alignment Scores:					
Pred. No.:	5,766-45	Length:	1104		
Score:	632.00	Matches:	121		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	12	Gaps:	0		
US-10-001-885-125 (1-121) x BG311398 (1-1104)					
QY	1 MetValaNgylleLeuAlaAsmGylGluIleValGlnAspAspProArgValaArgThr	20			
Db	155 ATGGTCCGATCTTGGCCATGGGGAATGTCAAGATACAGACCCCGAGTAGACC	211			
QY	21 ThrThrglnProProArgGlySerIleProArgGlnSerPhePheAsnArgGlyHisIy	40			
Db	212 ACTACCCAGCCACCAAGAGGTAGCATTCCTCGACAGAGTTTTCATATAGGGGCCATG	271			
QY	41 AlaProProGlyGlyProGlyProArgGlnGlnGlnAlaGlyAlaArgLeuGlyAlaAla	60			
Db	272 GGTCCCCAGAGGAGCTCTGGCCCCCGACAGACAGACAGAGGCCAGGCTGGGTGCTCT	331			

Oy	61	GlnSerProPheAsnAspLeuSerAsnArgIleuValLeuMetGlyPheProGlnTrpHis	80
Db	332	CAATCCCCCTTCAGATGACCTCAACCGGACAGCTGGTGAACATGGGCTTCCGCACTGGCA	391
Oy	81	LeuGlyAsnHisAlaValGluProValThirSerIleLeuLeuPheLeuMetMet	100
Db	392	CTGGCAACCACTGCTGTGGAGCGGGTGAACCTCAATCCGCTCCTTCCGTGCTCAGATG	451
Oy	101	LeuGlyValArgGlyLeuLeuLeuValGlyLeuValTyrIleuValSerHisLeuSerGln	120
Db	452	CTTGGTGTTCGTGGGCTCTCTCTGTGGCCTTGTCTTCACTGGTGGTCCACCTGAGTACG	511
Oy	121	Arg	121
Db	512	CGG	514

[illegible]

REFERENCE	1 (bases 1 to 528)
AUTHORS	Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL	unpublished (2000)
COMMENT	Contact: John Quackenbush

The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: [johnge@igr.org](mailto:johnge@igr.org)  
Place: 148  
Seq primer: Reverse.

```
FEATURES
source
location/Qualifiers
1..528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_1b="MAGE resequences, MAGE"
/note="Vector: pBluescriptSkM"
```

	Alignment Scores:			
	Pred. No.:	3,456-44	Length:	528
Score:	618.00	Matches:		118
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	97.78%	Indels:	0	
DB:	10	Gaps:	0	
 US-10-001-885-125 (1-121) X AW960624 (1-528)				
QY	4 TleLeuAlaaenGlyGlnIleValGIAnspAspProArgVaLaagThrThrInGln	23		
Dd	3 AACTTGGCCMAAGGGGAATCGTGAGGATATACACCCCCAGAGGAGACCCTACCCAG	62		
QY	24 ProPArGlySerIleProArgGlnSerPhePheAsnArgLYIHsiGLyalaProPro	43		
Dd	63 CCACCAAGAAGTAGATTCTCTGCACAGACTTTCTTAACAAGGGGCCATGTGTCTCCCCA	122		
QY	44 GLyGLyProGLyProArgGlnGlnGlnAlaGLyalaLaargLeuGLyalaLaaglInsPro	63		
Dd	123 GGAGGATCCTGGGCCCCCGCACACAGCAGTCACGCTGGGTGTCTCATGTCCTCCC	182		